Complete Approximation of Horn DL Ontologies

Anneke Bötcher[1], Carsten Lutz[1], and Frank Wolter[2]

[1] University of Bremen, Germany
anneke@uni-bremen.de clu@uni-bremen.de
[2] University of Liverpool, United Kingdom
wolter@liverpool.ac.uk

Abstract. We study the approximation of expressive Horn DL ontologies in less expressive Horn DLs, with completeness guarantees. Cases of interest include Horn-$\mathcal{SRIF}$-to-$\mathcal{ELR}_\bot$, Horn-$\mathcal{SHIF}$-to-$\mathcal{ELH}_\bot$, and others. Since finite approximations almost never exist, we carefully map out the structure of infinite approximations. This provides a solid theoretical foundation for constructing incomplete approximations in practice in a controlled way. Technically, we exibit a connection to the axiomatization of quasi-equations valid in classes of semilattices with operators and additionally develop a direct proof strategy based on the chase and on homomorphisms that allows us to also deal with approximations of bounded role depth.

1 Introduction

There is a large number of description logics that vary considerably regarding their expressive power and computational properties [2] and despite prominent standardization efforts, many different DLs continue to be used in ontologies from practical applications.[3] As a result, it is often necessary to convert an ontology formulated in some DL $\mathcal{L}$ into another DL $\mathcal{L}'$, a particularly important case being that $\mathcal{L}'$ is a fragment of $\mathcal{L}$—if it is not, then one could use the fragment $\mathcal{L} \cap \mathcal{L}'$ of $\mathcal{L}$ in place of $\mathcal{L}'$. For example, this happens in *ontology import* when an engineer who designs an ontology formulated in $\mathcal{L}'$ wants to reuse content from an existing ontology formulated in $\mathcal{L}$. The problem that emerges from this is *ontology approximation*, a form of knowledge compilation [27, 11].

In this paper, we are interested in approximating an ontology $\mathcal{O}_E$ formulated in an expressive description logic $\mathcal{L}$ by an ontology $\mathcal{O}_L$ formulated in a fragment $\mathcal{L}'$ of $\mathcal{L}$. We aim to construct $\mathcal{O}_L$ so that it preserves all information from $\mathcal{O}_E$ expressible in $\mathcal{L}'$, called a *greatest lower bound* in knowledge compilation [27]. Formally, for every $\mathcal{L}'$ concept inclusion $C \sqsubseteq D$ that is formulated in the signature Σ of $\mathcal{O}_E$, we require that $\mathcal{O}_E \models C \sqsubseteq D$ if and only if $\mathcal{O}_L \models C \sqsubseteq D$, and likewise for role inclusions and any other type of ontology statement supported by $\mathcal{L}'$. We say that $\mathcal{O}_L$ is *sound* as an approximation if it satisfies the "if" part of this property and *complete* if it satisfies the "only if" part. It is equivalent to

[3] See, for example, the BioPortal repository at https://bioportal.bioontology.org/.

demand that $\mathcal{O}_E \models \mathcal{O}_L$ and for every $\mathcal{L}$ ontology $\mathcal{O}$ in signature Σ, $\mathcal{O}_E \models \mathcal{O}$ implies $\mathcal{O}_L \models \mathcal{O}$. We consider both the case where $\mathcal{O}_L$ is formulated in Σ (*non-projective* approximation) and the case where additional symbols are admitted (*projective* approximation).

In practice, approximations are often constructed in an ad hoc way that is sound but not complete. For example, it is common to simply replace all subconcepts of $\mathcal{O}_E$ that use a constructor which is not available in $\mathcal{L}'$ with top or with bottom. In fact, full completeness is typically not attainable as it brings about infinite ontologies even in simple cases. For example, consider the $\mathcal{ELI}$ ontology $\mathcal{O}_E = \{\exists \mathtt{hasSupervisor}^-.\top \sqsubseteq \mathtt{Manager}\}$ which contains only a single range restriction. There is no finite $\mathcal{EL}$ approximation $\mathcal{O}_L$ since for all $n \geq 1$, $\mathcal{O}_E$ entails the $\mathcal{EL}$ concept inclusion

$$\exists \mathtt{hasSupervisor}^n.\top \sqsubseteq \exists \mathtt{hasSupervisor}.(\mathtt{Manager} \sqcap \exists \mathtt{hasSupervisor}^{n-1}.\top).$$

Also for the $\mathcal{ELF}$ ontology $\mathcal{O}_E = \{\mathtt{func}(\mathtt{hasSupervisor}), \mathtt{func}(\mathtt{reportsTo})\}$, there is no finite $\mathcal{EL}$ approximation $\mathcal{O}_L$ since for all $n, m \geq 1$, $\mathcal{O}_E$ entails the $\mathcal{EL}$ concept inclusion[4]

$$\begin{aligned}\exists \mathtt{reportsTo}.\exists \mathtt{hasSupervisor}^n.\top \sqcap \exists \mathtt{reportsTo}^m.\top \ \sqsubseteq \\ \exists \mathtt{reportsTo}.(\exists \mathtt{hasSupervisor}^n.\top \sqcap \exists \mathtt{reportsTo}^{m-1}.\top).\end{aligned}$$

Nevertheless, ad hoc ways to construct approximations can be a problematic choice since it is often only poorly understood which information has been given up. This is particularly worrying in ontology import where one would like to construct an as meaningful approximation as possible, rather than just *some* approximation. Our aim is to address this problem by carefully studying the structure of infinite approximations. As the expressive DL $\mathcal{L}$, we consider Horn-$\mathcal{SRIF}$ and fragments thereof. As the fragment $\mathcal{L}'$, we consider $\mathcal{ELR}_\bot$ and corresponding fragments thereof, where $\mathcal{ELR}_\bot$ denotes the extension of $\mathcal{ELH}_\bot$ with role inclusions of the form $r_1 \circ \cdots \circ r_n \sqsubseteq r$. For example, we study Horn-$\mathcal{SRIF}$-to-$\mathcal{ELR}_\bot$ approximation, Horn-$\mathcal{SHIF}$-to-$\mathcal{ELH}_\bot$, $\mathcal{ELI}$-to-$\mathcal{EL}$, $\mathcal{ELF}$-to-$\mathcal{EL}$, and so on. Subsumption is ExpTime-complete in all mentioned source DLs and PTime-complete in all mentioned target DLs [1, 20]. We thus support ontology designers who build an ontology in a tractable DL and want to import in a well-understood way from an existing ontology formulated in a more costly DL, without compromising tractability.

We provide the following results. In Section 3, we present non-projective approximations for the $\mathcal{ELR}^\mathcal{I}\mathcal{F}_\bot$-to-$\mathcal{ELR}_\bot$ case and several subcases such as $\mathcal{ELH}^\mathcal{I}_\bot$-to-$\mathcal{ELH}_\bot$. Here, superscript $\mathcal{I}$ means that inverse roles are admitted only in role hierarchies of the form $r \sqsubseteq s^-$ but not in concept inclusions and more complex role inclusions, which is actually a very common way to use inverse roles in practice. The presented approximation requires that $\mathcal{O}_E$ is *inverse closed*, meaning that for every role name r in $\mathcal{O}_E$, there is a role name $\hat{r}$ that is defined via

[4] A single functionality assertion would also do, but it is convenient for the example to have two role names in the signature of $\mathcal{O}_E$.

role hierarchies to be the inverse of r. This also yields *projective* approximations for the case where inverse closedness is not assumed and for the Horn-$\mathcal{SRIF}$-to-$\mathcal{ELR}_\bot$ case through a well-known normalization procedure. The completeness proof is based on a novel connection between ontology approximation and the axiomatizations of quasi-equations valid in classes of semilattices with operators (SLOs); note that SLOs have been used before to obtain algorithms for subsumption in extensions of $\mathcal{EL}$ [28, 29].

In Section 4, we construct non-projective $\mathcal{ELH}^{\mathcal{I}}\mathcal{F}_\bot$-to-$\mathcal{ELH}_\bot$ approximations under the mild assumption that whenever $\mathcal{O}_E \models r \sqsubseteq s^-$, then neither $\mathtt{func}(s) \in \mathcal{O}_E$ nor $\mathtt{func}(s^-) \in \mathcal{O}_E$. This again encompasses relevant subcases such as $\mathcal{ELHF}$-to-$\mathcal{ELH}$, without any syntactic assumptions. The main difference to Section 3 is how completeness is established. Here, we find a chase procedure for $\mathcal{L}$ ontologies that is inspired by and closely linked to the proposed approximation scheme, prove that it is complete for consequences formulated in $\mathcal{L}'$, and then show that consequences computed by the chase can also be derived using $\mathcal{O}_L$. In contrast to the approach based on SLOs, this enables us to also prove that finite approxmations always exist if we restrict the role depth of concepts in consequences to be bounded by a constant; we speak of *depth bounded approximations*.

We then proceed to study $\mathcal{ELI}_\bot$-to-$\mathcal{EL}_\bot$ approximations in Section 5. In contrast to the cases considered before, where both $\mathcal{L}$ and $\mathcal{L}'$ are based on the concept language $\mathcal{EL}_\bot$, here the *concept language* of $\mathcal{L}$ (which is $\mathcal{ELI}_\bot$) different from the one of $\mathcal{L}'$ (which is $\mathcal{EL}_\bot$). We present non-projective approximations for unrestricted ontologies $\mathcal{O}_E$ and for ontologies $\mathcal{O}_E$ which are in the well-known normal form for $\mathcal{ELI}_\bot$ ontologies that avoids syntactic nesting of concepts. The two approximation schemes are remarkably different, the latter arguably being more informative than the former.

In Section 6, we complement the proposed infinite approximations by showing that finite approximations do not exist even in simple cases and that depth bounded approximations are non-elementary in size even in simple cases.

Proof details are available in the appendix of the long version, available at `http://www.informatik.uni-bremen.de/tdki/research/papers.html`

Related Work. Approximation in a DL context was first studied in [27] where $\mathcal{FL}$ concepts are approximated by $\mathcal{FL}^-$ concepts and in [9] where $\mathcal{ALC}$ concepts are approximated by $\mathcal{ALE}$ concepts. In both cases, the approximation always exists, but ontologies are not considered. An incomplete approach to approximating $\mathcal{SHOIN}$ ontologies in DL-Lite$_\mathcal{F}$ is presented in [24] and complete (projective) approximations of $\mathcal{SROIQ}$ ontologies in DL-Lite$_\mathcal{A}$ are given in [7]. Such approximations are guaranteed to exist due to the limited expressive power of DL-Lite$_\mathcal{A}$. Approximation of Horn-$\mathcal{ALCHIQ}$ ontologies by DL-Lite$_\mathcal{R}$ ontologies in an OBDA context was considered in [8], exploiting the mapping formalism available in OBDA. In [22], approximation of $\mathcal{ELU}$ ontologies in terms of $\mathcal{EL}$ ontologies is studied, the main result being that it is ExpTime-hard and in 2ExpTime to decide a finite complete approximation exists. An incomplete approach to approximating $\mathcal{SROIQ}$ ontologies in $\mathcal{EL}^{++}$ is in [25]. There are also approaches towards efficient DL reasoning that involve computing approximations, which may be greatest lower bounds and/or least upper bounds. Such approximations are intentionally incomplete in order to not compromise efficiency, see for example [26, 14, 10]. Related to approximation is the problem whether a given $\mathcal{L}$ ontology can be equivalently rewritten into the fragment $\mathcal{L}'$ of $\mathcal{L}$, either non-projectively [21] or projectively [19]; note that this asks whether we have to approximate at all. There are also various approaches to OBDA with expressive DLs that involve forms of approximation such as [30, 31, 12, 15].

2 Preliminaries

Let $\mathsf{N_C}$ and $\mathsf{N_R}$ be disjoint and countably infinite sets of *concept* and *role names*. A *role* is a role name r or an *inverse role* r^-, with r a role name. A *Horn-$\mathcal{SRIF}$ concept inclusion (CI)* is of the form $L \sqsubseteq R$, where L and R are concepts defined by the syntax rules

$$R, R' ::= \top \mid \bot \mid A \mid \neg A \mid R \sqcap R' \mid \neg L \sqcup R \mid \exists \rho.R \mid \forall \rho.R$$
$$L, L' ::= \top \mid \bot \mid A \mid L \sqcap L' \mid L \sqcup L' \mid \exists \rho.L$$

with A ranging over concept names and ρ over roles. The *depth* of a concept R or L is the nesting depth of existential and universal restrictions in it. For example, the depth of $\exists r.B \sqcap \exists r.\exists s.A$ is two. A *Horn-$\mathcal{SRIF}$ ontology* $\mathcal{O}$ is a set of Horn-$\mathcal{SRIF}$ CIs, *functionality assertions* $\mathsf{func}(\rho)$, *transitivity assertions* $\mathsf{trans}(\rho)$, and *role inclusions (RIs)* $\rho_1 \circ \cdots \circ \rho_n \sqsubseteq \rho$. A role inclusion of the special form $\rho_1 \sqsubseteq \rho_2$ is a *role hierarchy (RH)*. We adopt the standard assumption that for any RI $\rho_1 \circ \cdots \circ \rho_n \sqsubseteq \rho$ with $n \geq 2$, we neither have $\mathcal{O} \models \mathtt{func}(\rho)$ nor $\mathcal{O} \models \mathtt{func}(\rho^-)$. Ontologies used in practice of course have to be finite. In this paper, though, we shall frequently consider also infinite ontologies.

An $\mathcal{ELRIF}_\bot$ *ontology* is a *Horn-$\mathcal{SRIF}$ ontology* in which both the left- and right-hand sides of CIs are $\mathcal{ELI}_\bot$ *concepts*, defined as usual, and that does not contain transitivity assertions. If $\mathcal{O}$ uses neither inverse roles nor functionality assertions, then it is an $\mathcal{ELR}_\bot$ *ontology*. We assume w.l.o.g. that, in $\mathcal{ELRIF}_\bot$ and below, the $\bot$ concept occurs only in the form $C \sqsubseteq \bot$. We assume that other standard DL names such as Horn-$\mathcal{SHIF}$ are understood. It should also be clear what we mean by saying that an ontology language is *based* on a concept language. For example, the ontology language $\mathcal{ELHIF}_\bot$ is based on the concept language $\mathcal{ELI}_\bot$. We also use a non-standard naming scheme, namely that $\mathcal{H}^\mathcal{I}$ indicates the presence of role hierarchies and of inverse roles in role hierarchies, but not in concept inclusions. $R^\mathcal{I}$ is similar, but additionally admits role inclusions while still restricting the use of inverse roles to role hierarchies.

For the semantics and more details on the relevant DLs, we refer to [2]. A *signature* Σ is a set of concept and role names. When speaking of $\mathcal{EL}(\Sigma)$ concepts, we mean $\mathcal{EL}$ concepts that only use concept and role names from Σ, and likewise for other concept languages. We use $\mathtt{sig}(\mathcal{O})$ to denote the set of concept and role names used in ontology $\mathcal{O}$.

Definition 1. *Let $\mathcal{O}_E$ be a Horn-$\mathcal{SRIF}$ ontology with sig$(\mathcal{O}_E) = \Sigma$ and let $\mathcal{L}$ be a fragment of Horn-$\mathcal{SRIF}$ based on the concept language $\mathcal{C}$. A (potentially infinite) $\mathcal{L}$ ontology $\mathcal{O}_L$ is an $\mathcal{L}$* approximation *of $\mathcal{O}_E$ if the following conditions are satisfied:*

1. *$\mathcal{O}_E \models C \sqsubseteq D$ iff $\mathcal{O}_L \models C \sqsubseteq D$ for all $\mathcal{C}(\Sigma)$ concepts C, D;*
2. *if α is a role inclusion, role hierarchy, or functionality assertion that falls within $\mathcal{L}$ and uses only symbols from Σ, then $\mathcal{O}_E \models \alpha$ iff $\mathcal{O}_L \models \alpha$.*

We call $\mathcal{O}_L$ a non-projective *$\mathcal{EL}$ approximation if $\mathtt{sig}(\mathcal{O}_L) \subseteq \Sigma$ and* projective *otherwise.*

*For $\ell \in \mathbb{N} \cup \{\omega\}$, (non-projective and projective) ℓ-*bounded $\mathcal{L}$ approximations *are defined analogously, except that only concepts C, D of depth bounded by ℓ are considered in Point 1.*

The term ω-bounded approximation, which is in fact the same as an unbounded approximation, is used only for uniformity. In Definition 1 and throughout the paper, we use $\mathcal{O}_E$ to denote ontologies formulated in an *E*xpressive DL and $\mathcal{O}_L$ to denote ontologies formulated in a *L*ightweight (in the sense of 'inexpressive') DL. Note that, trivially, infinite (non-projective and projective) approximations always exist: simply take as $\mathcal{O}_L$ the set of all relevant inclusions and assertions that are entailed by $\mathcal{O}_E$. Definition 1 speaks about Horn-$\mathcal{SRIF}$ ontologies $\mathcal{O}_E$ because this is the most expressive DL considered in this paper. In general, we speak about *$\mathcal{L}_S$-to-$\mathcal{L}_T$ approximation*, $\mathcal{L}_S$ a DL and $\mathcal{L}_T$ a fragment thereof, to denote the task of approximating an $\mathcal{L}_S$ ontology in $\mathcal{L}_T$. We call $\mathcal{L}_S$ the *source DL* and $\mathcal{L}_T$ the *target DL*. One can show that there are $\mathcal{ELI}$ ontologies $\mathcal{O}_E$ that have a finite projective $\mathcal{EL}$ approximation, but no finite non-projective $\mathcal{EL}$ approximation. Details are in the appendix. An alternative definition of approximations is obained by considering in Point 1 *any* $\mathcal{C}$ concept for C and D rather than only concepts in signature Σ. We choose not to do that because then even in the 1-bounded case, finite approximations do not necessarily exist.

Example 1. Consider the $\mathcal{ELI}$ ontology $\mathcal{O}_E = \{\exists r^-.A \sqsubseteq B\}$. We have as a consequence $A \sqcap \exists r.X \sqsubseteq \exists r.(B \sqcap X)$ for each of the infinitely many concept names $X \in \mathsf{N_C}$. Thus, every (projective or non-projective) 1-bounded $\mathcal{EL}$ approximation of $\mathcal{O}_E$ must be infinite under the alternative definition of approximation.

We now make some basic observations regarding approximations. The proof is straightforward.

Lemma 1. *Let $\mathcal{O}_E$ be a Horn-$\mathcal{SRIF}$ ontology with sig$(\mathcal{O}_E) = \Sigma$ and $\mathcal{L}$ a fragment of Horn-$\mathcal{SRIF}$. Then*

1. *an $\mathcal{L}$ ontology $\mathcal{O}_L$ is an $\mathcal{L}$ approximation of $\mathcal{O}_E$ iff $\mathcal{O}_E \models \mathcal{O}_L$ and for every $\mathcal{L}$ ontology $\mathcal{O}$ with $\mathcal{O}_E \models \mathcal{O}$ and $\mathtt{sig}(\mathcal{O}) \subseteq \Sigma$, $\mathcal{O}_L \models \mathcal{O}$;*
2. *$\bigcup_{i \geq 0} \mathcal{O}_\ell$ is an $\mathcal{L}$ approximation of $\mathcal{O}_E$ if for all $\ell \geq 0$, $\mathcal{O}_\ell$ is an ℓ-bounded $\mathcal{L}$ approximation of $\mathcal{O}_E$; the same is true for projective $\mathcal{L}$ approximations provided that $\mathtt{sig}(\mathcal{O}_\ell) \cap \mathtt{sig}(\mathcal{O}_{\ell'}) \subseteq \Sigma$ when $\ell \neq \ell'$.*

Point 1 may be viewed as an alternative definition of (non-projective) approximations. Point 2 is important because it allows us to concentrate on bounded approximations in proofs and to then obtain results for unbounded approximations as a byproduct. The following is well-known, see for example [5].

Lemma 2. *Given a Horn-$\mathcal{SRIF}$ ontology $\mathcal{O}_E$ with $\mathtt{sig}(\mathcal{O}) = \Sigma$, one can construct in polynomial time an $\mathcal{ELRIF}_\bot$ ontology $\mathcal{O}'_E$ with $\mathtt{sig}(\mathcal{O}'_E) \supseteq \Sigma$ that entails the same Horn-$\mathcal{SRIF}(\Sigma)$ concept inclusions, role inclusions, and functionality assertions.*

Note that the construction of the ontology $\mathcal{O}'_E$ from Lemma 2 requires the introduction of fresh concept names. Still, every ℓ-bounded $\mathcal{L}$ approximation of $\mathcal{O}'$ is a projective ℓ-bounded $\mathcal{L}$ approximation of $\mathcal{O}$. From now on, we work with $\mathcal{ELRIF}_\bot$ ontologies rather than with Horn-$\mathcal{SRIF}$ and thus obtain projective approximations also for the latter. Studying non-projective approximations of Horn-$\mathcal{SRIF}$ ontologies is outside the scope of this paper.

3 Unbounded $\mathcal{ELRIF}_\bot$-to-$\mathcal{ELR}_\bot$ Approximation

We provide (unbounded) approximations of $\mathcal{ELRIF}_\bot$ ontologies in $\mathcal{ELR}_\bot$. We assume throughout this section that $\mathcal{ELRIF}_\bot$ ontologies $\mathcal{O}_E$ are *inverse closed*, that is, for every role name r used in $\mathcal{O}_E$, there is a role name, which we denote $\hat{r}$, such that $r \sqsubseteq \hat{r}^-$ and $\hat{r} \sqsubseteq r^-$ are in $\mathcal{O}_E$. Thus, $\hat{r}$ is an *explicit name* for the inverse of r. We can clearly additionally assume w.l.o.g. that there are no other occurrences of inverse roles in $\mathcal{O}_E$, which we shall always do. In other words, it suffices to consider inverse closed $\mathcal{ELR}^\mathcal{I}\mathcal{F}_\bot$ ontologies in place of inverse closed $\mathcal{ELRIF}_\bot$ ontologies. We obtain non-projective approximations under this assumption, which clearly also yields projective approximations in the general case. The following theorem summarizes the results from this section.

Theorem 1. *Let $\mathcal{O}_E$ be an inverse closed $\mathcal{ELR}^\mathcal{I}\mathcal{F}_\bot$ ontology and $\Sigma = \mathsf{sig}(\mathcal{O}_E)$. Define $\mathcal{O}_L$ to be the $\mathcal{ELR}_\bot$ ontology that contains for all $\mathcal{EL}(\Sigma)$ concepts C, D and role names $r, s \in \Sigma$:*

1. *all CIs in $\mathcal{O}_E$;*
2. *$r \sqsubseteq s$ if $\mathcal{O}_E \models r \sqsubseteq s$;*
3. *$r_1 \circ \cdots \circ r_n \sqsubseteq r, \hat{r}_n \circ \cdots \circ \hat{r}_1 \sqsubseteq \hat{r}$ if $r_1 \circ \cdots \circ r_n \sqsubseteq r \in \mathcal{O}_E$ with $n \geq 2$;*
4. *$C \sqcap \exists r.D \sqsubseteq \exists r.(D \sqcap \exists \hat{r}.C)$;*
5. *$\exists r.C \sqcap \exists r.D \sqsubseteq \exists r.(C \sqcap D)$ if $\mathsf{func}(r) \in \mathcal{O}_E$;*
6. *$\exists r.\exists \hat{r}.C \sqsubseteq C$ if $\mathsf{func}(\hat{r}) \in \mathcal{O}_E$.*

Then $\mathcal{O}_L$ is an $\mathcal{ELR}_\bot$ approximation of $\mathcal{O}_E$.

Note that Points 1 to 3 essentially take over the part of $\mathcal{O}_E$ that is expressible in $\mathcal{ELR}_\bot$, Point 4 aims at capturing the consequences of inverse roles, Point 5 at functional roles, and Point 6 at the interaction between functional roles and inverse roles. Points 4 to 6 all introduce infinitely many CIs. The following example shows that Point 6 cannot be omitted.

Example 2. Consider $\mathcal{O}_E = \{\mathsf{func}(\hat{r}), r \sqsubseteq \hat{r}^-, \hat{r} \sqsubseteq r^-, A \sqsubseteq A\}$ Then $\mathcal{O}_E \models \exists r.\exists \hat{r}.A \sqsubseteq A$ but $\mathcal{O}'_L \not\models \exists r.\exists \hat{r}.A \sqsubseteq A$ for the ontology $\mathcal{O}'_L$ obtained from $\mathcal{O}_L$ by omitting the CIs of Point 6. To show this, consider the interpretation $\mathcal{I}$ with domain $\{0, 1, \ldots\}$ and

$$r^{\mathcal{I}} = \{(2n, 2n+1) \mid n \geq 0\}, \hat{r}^{\mathcal{I}} = \{(2n+1, 2n+2) \mid n \geq 0\}, A^{\mathcal{I}} = \{2n \mid n \geq 1\}$$

Then $\mathcal{I}$ is a model of $\mathcal{O}'_L$ but $0 \in (\exists r.\exists \hat{r}.A)^{\mathcal{I}} \setminus A^{\mathcal{I}}$.

It should be obvious how Point 5 captures the $\mathcal{ELF}$-to-$\mathcal{EL}$ example from the introduction. Point 4 captures the natural variation of the $\mathcal{ELI}$-to-$\mathcal{EL}$ example from the introduction obtained by converting the $\mathcal{ELI}$ ontology $\mathcal{O}_E$ used there into an inverse closed $\mathcal{ELH}^{\mathcal{I}}$ ontology, as follows.

Example 3. Let $\mathcal{O}_E = \{\exists \mathtt{supervises}.\top \sqsubseteq \mathtt{Manager}, \mathtt{hasSuper} \sqsubseteq \mathtt{supervises}^-, \mathtt{supervises}^- \sqsubseteq \mathtt{hasSuper}\}$. Point 4 yields

$$\exists \mathtt{hasSuper}^n.\top \sqsubseteq \exists \mathtt{hasSuper}.(\exists \mathtt{supervises}.\top \sqcap \exists \mathtt{hasSuper}^{n-1}.\top)$$

which together with $\exists \mathtt{supervises}.\top \sqsubseteq \mathtt{Manager} \in \mathcal{O}_L$ yields the desired

$$\exists \mathtt{hasSuper}^n.\top \sqsubseteq \exists \mathtt{hasSuper}.(\mathtt{Manager} \sqcap \exists \mathtt{hasSuper}^{n-1}.\top).$$

Theorem 1 also settles several natural subcases of (projective) $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\bot$-to-$\mathcal{ELR}_\bot$ approximation such as $\mathcal{ELH}^{\mathcal{I}}$-to-$\mathcal{ELH}$. For subcases where the source DL does not contain inverse roles such as $\mathcal{ELF}$-to-$\mathcal{EL}$, the concept inclusions in Point 4 are still present in the approximation as we still assume inverse closedness. This could also be avoided, as in the results presented in the subsequent section. We find it remarkable that the construction of $\mathcal{O}_L$ is based almost entirely on a purely syntactic analysis of $\mathcal{O}_E$, rather than involving reasoning. Reasoning is only required to derive the role hierarchies to be included in $\mathcal{O}_L$, in Point 2 of Theorem 1. Although we do not consider this aspect very important, we mention that this problem is ExpTime-complete when $\mathcal{O}_E$ is formulated in $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\bot$ and in PTime for many of the captures subcases such as when $\mathcal{O}_E$ is formulated in $\mathcal{ELR}^{\mathcal{I}}$.

It is straightforward to show that the ontology $\mathcal{O}_L$ from Theorem 1 is sound as an approximation. To prove completeness, we establish a novel connection between $\mathcal{EL}_\bot$ approximations and axiomatizations of the quasi-equations that are valid in classes of semilattices with operators (SLOs) [16, 29, 17]. Roughly speaking, an approximation is obtained from such an axiomatization by instantiating its equations, which correspond (in the sense of modal correspondence theory) to the role inclusions and hierarchies in the original ontology, with arbitrary $\mathcal{EL}$ concepts. A detailed presentation is provided in the appendix.

4 Depth-Bounded $\mathcal{ELH}^{\mathcal{I}}\mathcal{F}_\bot$-to-$\mathcal{ELH}_\bot$ Approximation

We pursue an alternative approach to proving that an approximation is complete, based on a suitable version of the chase. This allows us to also treat the case of depth bounded approximations. Moreover, approximations obtained in this section are non-projective and the assumption of ontologies being inverse closed is not needed. We consider $\mathcal{ELH}^{\mathcal{I}}\mathcal{F}_\bot$-to-$\mathcal{ELH}_\bot$ approximation and subcases thereof. An extension to role inclusions should be possible, but is left as future work.

We assume w.l.o.g. that role hierarchies only take the two forms $r \sqsubseteq s$ and $r \sqsubseteq s^-$. We further assume the following syntactic restriction:

($\heartsuit$) $\mathcal{O}_E \models r \sqsubseteq s^-$ implies that neither $\mathtt{func}(s) \in \mathcal{O}_E$ nor $\mathtt{func}(s^-) \in \mathcal{O}_E$.

This assumption is not without loss of generality, it serves to eliminate a subtle interaction between inverse roles, functional roles, and role hierarchies. While this interaction is captured implicitly and gracefully by the unbounded approximation scheme given in the previous section, it causes complications in the depth bounded case. We briefly comment on this at the end of the section.

Let C be an $\mathcal{EL}_\bot$ concept and $k \geq 0$. By *decorating C with subconcepts from $\mathcal{O}_E$ at leaves*, we mean to replace any number of occurrences of a quantifier-free subconcept D by a concept $D \sqcap D_1 \sqcap \cdots \sqcap D_k$, $D_1, \ldots, D_k$ subconcepts of $\mathcal{O}_E$. The following is the main result of this section.

Theorem 2. *Let $\mathcal{O}_E$ be an $\mathcal{ELFH}^{\mathcal{I}}_\bot$ ontology, $\Sigma = \mathtt{sig}(\mathcal{O}_E)$, and $\ell \in \mathbb{N} \cup \{\omega\}$ a depth bound. Define $\mathcal{O}_L$ to be the $\mathcal{ELH}_\bot$ ontology that consists of the following, where $\ell' = \max\{0, \ell - 1\}$ and r, r_1, r_2, s are role names from Σ:*

1. *all concept inclusions from $\mathcal{O}_E$;*
2. *$r \sqsubseteq s$ if $\mathcal{O}_E \models r \sqsubseteq s$;*
3. *$C_1 \sqcap \exists r.C_2 \sqsubseteq \exists r.(C_2 \sqcap \exists s.C_1)$ if $\mathcal{O}_E \models r \sqsubseteq s^-$, $\exists s.C_1$ is a subconcept of $\mathcal{O}_E$ or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ, and C_2 is an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ' decorated with subconcepts of $\mathcal{O}_E$ at leaves;*
4. *$\exists r_1.C_1 \sqcap \exists r_2.C_2 \sqsubseteq \exists r_1.(C_1 \sqcap C_2)$ if there is a role name s with $\mathcal{O}_E \models r_1 \sqsubseteq s$, $\mathcal{O}_E \models r_2 \sqsubseteq s$, and $\mathcal{O}_E \ni \mathtt{func}(s)$, and C_1, C_2 are $\mathcal{EL}(\Sigma)$ concepts of depth bounded by ℓ' decorated with subconcepts of $\mathcal{O}_E$ at leaves.*

Then $\mathcal{O}_L$ is an ℓ-bounded approximation of $\mathcal{O}_E$.

We tried to be as economic as possible regarding the classes of concepts that have to be considered in Points 3 and 4, which has led to the subtle depth bounds stated there. Note that Theorem 2 also settles the cases of $\mathcal{ELFH}_\bot$-to-$\mathcal{ELH}_\bot$ approximation and of $\mathcal{ELH}^{\mathcal{I}}_\bot$-to-$\mathcal{ELH}_\bot$ approximation (both without any syntactic restrictions), as well as the variation of all these cases without $\mathcal{H}$ and/or $\bot$ in both the source and target DL.

Due to Points 3 and 4, the approximation $\mathcal{O}_L$ is of (single) exponential size even when $\ell = 0$. This must necessarily be the case because otherwise we would obtain a subexponential algorithm for the ExpTime-complete subsumption problem between concept names in $\mathcal{ELI}$. Note that Theorem 2 also reproves the upper bound for this problem: compute the 0-bounded approximation of single exponential size in exponential time and then decide subsumption in $\mathcal{EL}$ in PTime.

It is again straightforward to verify that the ontology $\mathcal{O}_L$ constructed in Theorem 2 is sound as an approximation, that is, $\mathcal{O}_E \models \mathcal{O}_L$. Completeness is established in two steps. First, we introduce a suitable version of the chase and show that it is sound and complete regarding the consequences relevant for approximation, and second we show that the CIs in $\mathcal{O}_L$ can simulate derivations of this chase.

Let us now drop assumption ($\heartsuit$). One can prove that we then need to extend Points 1 to 4 of Theorem 2 with the following:

5. $\exists r.\exists s.C \sqsubseteq C$ if $\mathcal{O}_E \models r \sqsubseteq s^-$, $\mathtt{func}(s) \in \mathcal{O}_E$, and C is a subconcept of $\mathcal{O}_E$ or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ;
6. $\exists s.\exists r.C \sqsubseteq C$ if $\mathcal{O}_E \models r \sqsubseteq s^-$, $\mathtt{func}(s^-) \in \mathcal{O}_E$, and C is a subconcept of $\mathcal{O}_E$ or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ.

However, this is still not sufficient to obtain a complete approximation. Consider the $\mathcal{ELH}^\mathcal{I}$ ontology

$$\mathcal{O}_E = \{A \sqsubseteq \exists r_1.\exists r_2.(B \sqcap \exists s.\top), \quad s \sqsubseteq r_1, \quad s \sqsubseteq r_2^-, \quad \mathtt{func}(r_1^-), \quad \mathtt{func}(r_2^-)\}.$$

It can be verified that $\mathcal{O}_E \models A \sqsubseteq B$. However, it can also be proved that even when $\mathcal{O}_L$ is the set of all statements from Points 1 to 6 with $\ell = 0$, $\mathcal{O}_L \not\models A \sqsubseteq B$.

5 $\mathcal{ELI}_\bot$-to-$\mathcal{EL}_\bot$ Approximation

The previous sections concentrated on cases of $\mathcal{L}$-to-$\mathcal{L}'$ approximations where both $\mathcal{L}$ and $\mathcal{L}'$ were based on the same concept language $\mathcal{EL}_\bot$. In this section, we take a look at $\mathcal{ELI}_\bot$-to-$\mathcal{EL}_\bot$ approximation, thus aiming to approximate away inverse roles in concepts inclusions. We consider bounded and unbounded, projective and non-projective approximations. The proof techniques is based on the chase, as in the previous section.

Constructing informative non-projective approximations appears to be difficult in the $\mathcal{ELI}_\bot$-to-$\mathcal{EL}_\bot$ case. Informally, concepts of the form $\exists r^-.C$ can be used as a marker invisible to $\mathcal{EL}_\bot$ that is propagated along role edges, resulting in rather complex $\mathcal{EL}$ concept inclusions to be entailed by $\mathcal{O}_E$.

Example 4. Let $\mathcal{O}_E = \{A \sqsubseteq \exists s^-.\top, \exists r^-.\exists s^-.\top \sqsubseteq \exists s^-.\top, \exists s^-.\top \sqsubseteq B\}$. Then $\mathcal{O}_E \models C \sqsubseteq C'$ for all $\mathcal{EL}$ concepts C, C' where C' is obtained from C by decorating with B any node that is reachable in C from a node decorated with A along an r-path (we view an $\mathcal{EL}$ concept as a tree in the standard way, see for example [18]).

We now give a non-projective approximation that captures the effects demonstrated in Example 4. For an $\mathcal{ELI}_\bot$ ontology $\mathcal{O}_E$, we use $\mathtt{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ to denote the set of all $\mathcal{EL}$ concepts that can be obtained by starting with a subconcept of a concept from $\mathcal{O}_E$ and then replacing every subconcept of the form $\exists r^-.D$ with $\top$. Let C be an $\mathcal{EL}$ concept. An $\mathcal{EL}$ concept C' is a $\mathtt{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ *decoration of* C if it can be obtained from C by conjunctively adding concepts from $\mathtt{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ to a single occurrence of a subconcept in C.

Theorem 3. *Let $\mathcal{O}_E$ be an $\mathcal{ELI}_\bot$ ontology and $\Sigma = \mathtt{sig}(\mathcal{O}_E)$. Define an $\mathcal{EL}_\bot$ ontology $\mathcal{O}_L$ that consists of the following:*

1. *$C \sqsubseteq C'$ if $\mathcal{O}_E \models C \sqsubseteq C'$, C an $\mathcal{EL}(\Sigma)$ concept and C' a $\mathtt{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ decoration of C;*
2. *$C \sqsubseteq \bot$ if $\mathcal{O}_E \models C \sqsubseteq \bot$, C an $\mathcal{EL}(\Sigma)$ concept;*

Then $\mathcal{O}_L$ is an approximation of $\mathcal{O}_E$.

We prove completeness by using the standard chase for $\mathcal{ELI}_\bot$ that is complete also for $\mathcal{ELI}$ consequences and then work with homomorphisms that are in a certain sense 'forwards directed'. Details are given in the appendix.

Arguably, the approximation provided by Theorem 3 is less informative than the ones obtained in the previous sections. We next demonstrate that in the projective case, more informative approximations can be constructed. One way of doing this is to first convert the $\mathcal{ELI}_\bot$ ontology into an inverse closed $\mathcal{ELH}^{\mathcal{I}}_\bot$ ontology as in Section 3. Here, we pursue a natural alternative that consists of first converting the $\mathcal{ELI}_\bot$ ontology into a widely known normal form for such ontologies [2] and then providing non-projective approximations for ontologies in normal form. Note that, in practice, ontologies are sometimes already constructed in this normal form or at least in a form very close to it.

An $\mathcal{ELI}_\bot$ ontology $\mathcal{O}$ is in *normal form* if all CIs in $\mathcal{O}$ have one of the forms $\top \sqsubseteq A_1$, $A_1 \sqsubseteq \bot$, $A_1 \sqsubseteq \exists \rho.A_2$, $\exists \rho.A_1 \sqsubseteq B$, and $A_1 \sqcap \cdots \sqcap A_n \sqsubseteq B$ where $A_1, \ldots, A_n, B$ range over concept names and ρ ranges over roles. It is well known that every $\mathcal{ELI}_\bot$ ontology $\mathcal{O}$ can be converted in an $\mathcal{ELI}_\bot$ ontology $\mathcal{O}'$ in linear time such that $\mathcal{O}'$ is in normal form and a conservative extension of $\mathcal{O}$. Clearly, any (projective or non-projective) approximation of $\mathcal{O}'$ is a projective approximation of $\mathcal{O}$.

Theorem 4. *Let $\mathcal{O}_E$ be an $\mathcal{ELI}_\bot$ ontology in normal form, $\Sigma = \mathtt{sig}(\mathcal{O}_E)$, and $\ell \in \mathbb{N} \cup \{\omega\}$ a depth bound. Define an $\mathcal{EL}_\bot$ ontology $\mathcal{O}_L$ that consists of the following:*

1. *all concept inclusions from $\mathcal{O}_E$ that are of the form $\top \sqsubseteq A$, $A \sqsubseteq \bot$, $\exists r.A \sqsubseteq B$, or $A \sqsubseteq \exists r.B$,;*
2. *$A_1 \sqcap \cdots \sqcap A_n \sqsubseteq B$ if $\mathcal{O}_E \models A_1 \sqcap \cdots \sqcap A_n \sqsubseteq B$, $A_1, \ldots, A_n, B \in \mathsf{N_C}$ in $\mathcal{O}_E$;*
3. *$A \sqcap \exists r.C \sqsubseteq \exists r.(C \sqcap B)$ if $\exists r^-.A \sqsubseteq B \in \mathcal{O}_E$ and C is an $\mathcal{EL}(\Sigma)$ concept of depth bounded by $\ell - 1$.*

Then $\mathcal{O}_L$ is an appoximation of $\mathcal{O}_E$.

It is straightforward to verify that $\mathcal{O}_L$ is sound. To prove completeness, we use the same strategy as for Theorem 2.

6 Size of Approximations

We prove that finite approximations do not necessarily exist and that depth bounded approximations can be non-elementary in size. These results hold both for projective and non-projective approximations and for all combinations of source and target DL considered in this paper. The ontologies used to prove these results are simple and show that for the vast majority of ontologies that occur in practical applications, neither finite approximations nor depth bounded approximations of elementary size can be expected. We focus on the cases $\mathcal{ELIH}$-to-$\mathcal{ELH}$, $\mathcal{ELHF}$-to-$\mathcal{ELH}$, and $\mathcal{ELH}^{\mathcal{I}}$-to-$\mathcal{ELH}$, starting with the non-existence of finite approximations.

Theorem 5. *None of the ontologies*

$$\{\exists r^{-}.A \sqsubseteq B\}, \quad \{\mathsf{func}(r), A \sqsubseteq A\}, \quad \{r \sqsubseteq s^{-}, A \sqsubseteq A\}$$

has finite projective $\mathcal{ELH}$ approximations.

We next show that bounded depth approximations can be non-elementary in size. The function $\texttt{tower} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined as $\texttt{tower}(0, n) := n$ and $\texttt{tower}(k+1, n) := 2^{\texttt{tower}(k,n)}$. The *size* of a (finite) ontology is the number of symbols needed to write it, with concept and role names counting as one. We use Γ_n to denote a fixed finite tautological set of $\mathcal{EL}$ concept inclusions that contains the symbols $\Sigma_n = \{r_1, r_2, A_1, \hat{A}_1, \dots, A_n, \hat{A}_n\}$. One could take, for example, $\exists r_1.\top \sqsubseteq \top$, $\exists r_2.\top \sqsubseteq \top$, and all $A \sqsubseteq A$ with $A \in \{A_1, \hat{A}_1, \dots, A_n, \hat{A}_n\}$.

Theorem 6. *Let $n \geq 0$ and let $\mathcal{O}_n$ be the union of Γ_n with any of the following sets:*

$$\{\exists r^{-}.A \sqsubseteq B\}, \quad \{\mathsf{func}(r), A \sqsubseteq A\}, \quad \{r \sqsubseteq s^{-}, A \sqsubseteq A\}$$

For every $\ell \geq 1$, any ℓ-bounded projective $\mathcal{ELH}$ approximation $\mathcal{O}_L$ of $\mathcal{O}_n$ must be of size at least $\texttt{tower}(\ell, n)$.

7 Conclusion

There are several questions that emerge from our work. For example, it remains an open problem to develop a convincing approximation for $\mathcal{ELH}^{\mathcal{I}}\mathcal{F}_\bot$-to-$\mathcal{ELH}_\bot$ in the non-projective case, or even for Horn-$\mathcal{SRIF}$-to-$\mathcal{ELR}_\bot$. It would also be useful to consider more expressive target DLs such as the extension of $\mathcal{ELH}_\bot$ or $\mathcal{ELR}_\bot$ with range restrictions, and to add nominals to the picture. Of course, it would also be very interesting to approximate non-Horn DLs such as $\mathcal{ALC}$, $\mathcal{SHIQ}$, and $\mathcal{SROIQ}$ in (tractable and intractable) Horn DLs.

From a conceptual perspective, it would be of great interest to understand how approximations can be tailored towards intended applications. In this context, observe that all our bounded depth approximation schemes still work when the set of concepts of depth ℓ is replaced with any set Γ of concepts closed under subconcepts. For example, if one wants to decide subsumption between $\mathcal{EL}$ concepts C and D relative to an $\mathcal{ELI}$ ontology $\mathcal{O}_E$, one can approximate $\mathcal{O}_E$ in $\mathcal{EL}$ relative to the set Γ of subconcepts of C and D; the resulting $\mathcal{EL}$ ontology $\mathcal{O}_L$ will entail $C \sqsubseteq D$ iff $\mathcal{O}_E$ does. In a similar spirit, it would be interesting to develop approximations that aim at query answering applications.

Acknowledgments. Supported by the DFG Collaborative Research Center 1320 EASE - Everyday Activity Science and Engineering.

References

1. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. In: Proc. of IJCAI. Morgan Kaufmann (2005)
2. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
3. Beklemishev, L.D.: Positive provability logic for uniform reflection principles. Ann. Pure Appl. Logic **165**(1), 82–105 (2014)
4. van Benthem, J.: Correspondence theory. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, vol. 2, pp. 167–247. Reidel, Dordrecht (1984)
5. Bienvenu, M., Hansen, P., Lutz, C., Wolter, F.: First order-rewritability and containment of conjunctive queries in horn description logics. In: Proceedings of IJCAI. pp. 965–971 (2016)
6. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press (2001)
7. Botoeva, E., Calvanese, D., Rodriguez-Muro, M.: Expressive approximations in *DL-Lite* ontologies. In: AIMSA. pp. 21–31 (2010)
8. Botoeva, E., Calvanese, D., Santarelli, V., Savo, D.F., Solimando, A., Xiao, G.: Beyond OWL 2 QL in OBDA: rewritings and approximations. In: Proc. of AAAI. pp. 921–928. AAAI Press (2016)
9. Brandt, S., Küsters, R., Turhan, A.Y.: Approximation and difference in description logics. In: KR. pp. 203–214 (2002)
10. Carral, D., Feier, C., Grau, B.C., Hitzler, P., Horrocks, I.: *EL*-ifying ontologies. In: Proc. of IJCAR. pp. 464–479 (2014)
11. Darwiche, A., Marquis, P.: A knowledge compilation map. J. Artif. Intell. Res. **17**, 229–264 (2002)
12. Feier, C., Lutz, C., Kuusisto, A.: Rewritability in monadic disjunctive datalog, mmsnp, and expressive description logics. In: Proceedings of the 20th International Conference on Database Theory (ICDT17) (2017)
13. Goldblatt, R.: Varieties of complex algebras. Annals of Pure and Applied Logic **44**, 173–242 (1989)
14. Groot, P., Stuckenschmidt, H., Wache, H.: Approximating description logic classification for semantic web reasoning. In: Proc. of ESWC. pp. 318–332 (2005)
15. Hernich, A., Lutz, C., Papacchini, F., Wolter, F.: Horn-rewritability vs ptime query evaluation in ontology-mediated querying. In: Proceedings of IJCAI-ECAI. AAAI Press (2018)
16. Jackson, M.: Semilattices with closure. Algebra Universalis **52**, 1–37 (2004)
17. Kikot, S., Kurucz, Á., Tanaka, Y., Wolter, F., Zakharyaschev, M.: Kripke completeness of strictly positive modal logics over meet-semilattices with operators. CoRR **abs/1708.03403** (2017), `http://arxiv.org/abs/1708.03403`
18. Konev, B., Lutz, C., Ozaki, A., Wolter, F.: Exact learning of lightweight description logic ontologies. Journal of Machine Learning Research **18**(201), 1–63 (2018)
19. Konev, B., Lutz, C., Wolter, F., Zakharyaschev, M.: Conservative rewritability of description logic TBoxes. In: Proc. of IJCAI. pp. 1153–1159 (2016)
20. Krötzsch, M., Rudolph, S., Hitzler, P.: Complexities of horn description logics. ACM Trans. Comput. Log. **14**(1), 2:1–2:36 (2013)
21. Lutz, C., Piro, R., Wolter, F.: Description logic TBoxes: Model-theoretic characterizations and rewritability. In: IJCAI. pp. 983–988 (2011)
22. Lutz, C., Seylan, I., Wolter, F.: An automata-theoretic approach to uniform interpolation and approximation in the description logic $\mathcal{EL}$. In: Proc. of KR. AAAI Press (2012)

23. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic EL. J. Symb. Comput. **45**(2), 194–228 (2010)
24. Pan, J.Z., Thomas, E.: Approximating *OWL-DL* ontologies. In: AAAI. pp. 1434–1439 (2007)
25. Ren, Y., Pan, J.Z., Zhao, Y.: Soundness preserving approximation for TBox reasoning. In: AAAI (2010)
26. Schaerf, M., Cadoli, M.: Tractable reasoning via approximation. Artificial Intelligence **74**(2), 249 – 310 (1995)
27. Selman, B., Kautz, H.A.: Knowledge compilation and theory approximation. J. ACM **43**(2), 193–224 (1996)
28. Sofronie-Stokkermans, V.: Locality and subsumption testing in $\mathcal{EL}$ and some of its extensions. In: Areces, C., Goldblatt, R. (eds.) Advances in Modal Logic, vol. 7. pp. 315–339. College Publications (2008)
29. Sofronie-Stokkermans, V.: Representation theorems and locality for subsumption testing and interpolation in the description logics $\mathcal{EL}$ and $\mathcal{EL}^+$ and their extensions with n-ary roles and numerical domains. Fundam. Inform. **156**(3-4), 361–411 (2017)
30. Tserendorj, T., Rudolph, S., Krötzsch, M., Hitzler, P.: Approximate owl-reasoning with screech. In: Proc. of RR. LNCS, vol. 5341, pp. 165–180. Springer (2008)
31. Zhou, Y., Grau, B.C., Nenov, Y., Kaminski, M., Horrocks, I.: Pagoda: Pay-as-you-go ontology query answering using a datalog reasoner. J. Artif. Intell. Res. **54**, 309–367 (2015)

A Details for Section 2

We give an example of a $\mathcal{ELI}$ ontology with a finite projective $\mathcal{EL}$ approximation but no finite non-projective $\mathcal{EL}$ approximation.

Example 5. Consider the $\mathcal{ELI}$ ontology

$$\mathcal{O}_E = \mathcal{O} \cup \{A \sqsubseteq \exists r.\exists s^-.\top, \exists s^-.\top \sqsubseteq \exists r.\exists s^-.\top\}$$

where

$$\mathcal{O} = \{\exists s.\top \sqsubseteq \exists s.(A \sqcap X), X \sqsubseteq \exists r.(A \sqcap X), X \sqsubseteq \exists s.(A \sqcap X)\}.$$

Then an infinite non-projective $\mathcal{EL}$ approximation of $\mathcal{O}_E$ is $\mathcal{O}_L = \mathcal{O} \cup \{A \sqsubseteq \exists r^i.\top \mid i \geq 0\}$, there is no finite non-projective approximation, and a finite projective $\mathcal{EL}$ approximation for $\mathcal{O}_E$ is $\mathcal{O}_L = \mathcal{O} \cup \{A \sqsubseteq Y, Y \sqsubseteq \exists r.Y\}$.

B Details for Section 3

We show that one can obtain approximations from axiomatizations of the quasi-equations valid in classes of bounded semilattices with operators (SLOs). Theorem 1 is an instance of a general result stating that, under certain conditions, by identifying CIs with equations in the theory of SLOs, the substitution instances of the equations used in an axiomatization provide the additional CIs needed to approximate ontologies. This link between approximation and algebra can be used in a number of different ways: (1) existing axiomatization results can be used directly, as a black box, to obtain approximations; (2) if no axiomatization is available yet for the conditions on roles expressed in a DL of interest, the algebraic machinery can be used to determine a new axiomatization and, thereby, the corresponding approximation; (3) 'negative' results from algebra can be used to show that certain natural candidates for approximations do not work; (4) conversely, one can use approximation results to obtain axiomatizations of classes of SLOs. In fact, the direct approximation proofs presented in the next section provide a novel technique for obtaining axiomatizations of classes of SLOs. Note that the link to algebra does *not* provide any depth-bounded approximations from axiomatizations.

This section is structured as follows. After introducing the relevant algebraic notation, we prove a general result linking approximations to *complex* equational theories of SLOs, where an equational theory Ax of SLOs is complex if every SLO validating Ax can be respresented by subsets (complexes) of an interpretation validating Ax. This link is proved for equational theories of SLOs corresponding to arbitrary first-order conditions on roles. We then prove that any set P of functionality assertions and role inclusions that is inverse closed corresponds to a complex equational theory and apply this result to prove Theorem 1.

We introduce the relevant notation for semilattices with operators. A *bounded semilattice with monotone operators (SLO)* is an algebraic structure

$$\mathfrak{A} = (A, \wedge^{\mathfrak{A}}, \perp^{\mathfrak{A}}, \top^{\mathfrak{A}}, (\Diamond_r^{\mathfrak{A}} \mid r \in \mathcal{R}))$$

such that $(A, \wedge^{\mathfrak{A}}, \bot^{\mathfrak{A}}, \top^{\mathfrak{A}})$ is a *bounded semilattice* satisfying the equations

$$\forall x\ (x \wedge^{\mathfrak{A}} x \approx x) \tag{1}$$

$$\forall x\ \forall y\ (x \wedge^{\mathfrak{A}} y \approx y \wedge^{\mathfrak{A}} x) \tag{2}$$

$$\forall x\ \forall y\ \forall z\ (x \wedge^{\mathfrak{A}} (y \wedge^{\mathfrak{A}} z)) \approx (x \wedge^{\mathfrak{A}} y) \wedge^{\mathfrak{A}} z) \tag{3}$$

$$\forall x\ (x \wedge^{\mathfrak{A}} \top^{\mathfrak{A}} \approx x), \quad \forall x\ (x \wedge^{\mathfrak{A}} \bot^{\mathfrak{A}} \approx \bot^{\mathfrak{A}}) \tag{4}$$

and $\mathcal{R}$ is a set of role names such that the unary operators $\Diamond_r^{\mathfrak{A}}$, $r \in \mathcal{R}$, satisfy the equation

$$\forall x\ \forall y\ (\Diamond_r^{\mathfrak{A}}(x \wedge^{\mathfrak{A}} y) \wedge^{\mathfrak{A}} \Diamond_r^{\mathfrak{A}} y) \approx \Diamond_r^{\mathfrak{A}}(x \wedge^{\mathfrak{A}} y) \tag{5}$$

$$\Diamond_r^{\mathfrak{A}} \bot^{\mathfrak{A}} \approx \bot^{\mathfrak{A}} \tag{6}$$

In a SLO $\mathfrak{A}$, the partial order $\leq^{\mathfrak{A}}$ is defined as usual by taking $a \leq^{\mathfrak{A}} b$ iff $a \wedge^{\mathfrak{A}} b = a$, for all a, b in A. It is readily seen that $\Diamond_r^{\mathfrak{A}}$ is *monotone* with respect to $\leq^{\mathfrak{A}}$: if $a \leq^{\mathfrak{A}} b$ then $\Diamond_r^{\mathfrak{A}} a \leq^{\mathfrak{A}} \Diamond_r^{\mathfrak{A}} b$, for all a, b in A, and that $\Diamond_r^{\mathfrak{A}} \top^{\mathfrak{A}} = \top^{\mathfrak{A}}$. *SLO terms* τ over $\mathcal{R}$ are constructed from variables using the connectives $\wedge$, $\bot$, $\top$, and $\Diamond_r$, $r \in \mathcal{R}$, in the obvious way:

$$\tau, \sigma \quad := \quad x \ \mid \ \bot \ \mid \ \top \ \mid \tau \wedge \sigma \ \mid \Diamond_r \tau$$

where x ranges over a countably infinite set of variables. A *SLO equation* takes the form $\sigma \approx \tau$, where σ, τ are SLO terms; a *SLO quasi-equation* takes the form $\alpha_1 \wedge \cdots \wedge \alpha_n \to \alpha$, where $\alpha_1, \ldots, \alpha_n, \alpha$ are SLO equations. For SLO terms σ and τ we use $\sigma \leq \tau$ as a shorthand for the equation $\sigma \wedge \tau \approx \sigma$. A *valuation* v in a SLO $\mathfrak{A}$ is a mapping from the set of variables into A. The *value* $v(\tau)$ of a SLO term τ in $\mathfrak{A}$ is defined by induction over the construction of τ by setting $v(\bot) = \bot^{\mathfrak{A}}$, $v(\top) = \top^{\mathfrak{A}}$, $v(\sigma \wedge \tau) = v(\sigma) \wedge v(\tau)$, and $v(\Diamond_r \tau) = \Diamond_r^{\mathfrak{A}} v(\tau)$, for all role names $r \in \mathcal{R}$. An equation $\sigma \approx \tau$ is *true under* v *in* $\mathfrak{A}$ if $v(\sigma) = v(\tau)$. An equation $\sigma \approx \tau$ is *valid in* $\mathfrak{A}$, in symbols $\mathfrak{A} \models \sigma \approx \tau$, if $\sigma \approx \tau$ is true under all valuations in $\mathfrak{A}$. A quasi-equation $\rho = \alpha_1 \wedge \cdots \wedge \alpha_n \to \alpha$ is *valid in* $\mathfrak{A}$ if α is true under all valuations under which $\alpha_1, \ldots, \alpha_n$ are true. Equational theories of SLOs have been investigated in [16, 3, 29, 17].

Every SLO term τ defines an $\mathcal{EL}_\bot$ concept τ^C by replacing every variable x with a concept name A_x, and the connectives $\wedge$ and $\Diamond_r$ with $\sqcap$ and $\exists r$, respectively. For example, $(\Diamond_r x \wedge \top)^C = \exists r.A_x \sqcap \top$. Then any SLO equation $\alpha = (\sigma \leq \tau)$ defines the CI $\alpha^C = \sigma^C \sqsubseteq \tau^C$. We denote by $\cdot^T$ the obvious converse of $\cdot^S$ associating with every $\mathcal{EL}_\bot$ concept C (CI α) a SLO term C^T (SLO equation α^T, respectively). For example, $(\exists r.\exists r.A_x \sqsubseteq \exists r.A_x)^T = \Diamond_r \Diamond_r x \leq \Diamond_r x$.

We are in the position now to formulate the fundamental equivalence of $\mathcal{EL}_\bot$ TBox reasoning and the validity of SLO quasi-equations [29, 17].

Theorem 7. *For any $\mathcal{EL}_\bot$ ontology $\mathcal{O}$ and CI $C \sqsubseteq D$: $\mathcal{O} \models C \sqsubseteq D$ iff $\bigwedge_{\alpha \in \mathcal{O}} \alpha^T \to (C \sqsubseteq D)^T$ is valid in all SLOs.*

Theorem 7 and the correspondence between axiomatizations and approximations we are after are proved by showing that every SLO (validating a set Ax of equations) can be represented by the set of subsets of a DL interpretation (satisfying a role constraint P corresponding to Ax). In detail, every interpretation $\mathcal{I}$ defines a SLO $\mathcal{I}^+$ over any set $\mathcal{R}$ of role names by setting [13]:

$$\mathcal{I}^+ = (2^{\Delta^{\mathcal{I}}}, \wedge^{\mathcal{I}^+}, \bot^{\mathcal{I}^+}, \top^{\mathcal{I}^+}, (\Diamond_r^{\mathcal{I}^+} \mid r \in \mathcal{R})),$$

where for $X, Y \subseteq \Delta^{\mathcal{I}}$:

$$\begin{aligned} X \wedge^{\mathcal{I}^+} Y &= X \cap Y \\ \top^{\mathcal{I}^+} &= \Delta^{\mathcal{I}} \\ \bot^{\mathcal{I}^+} &= \emptyset \\ \Diamond_r^{\mathcal{I}^+} X &= \{d \in \Delta^{\mathcal{I}} \mid \exists d' \in X \ (d, d') \in r^{\mathcal{I}}\} \end{aligned}$$

Observe that the definition of the SLO $\mathcal{I}^+$ does not depend on the interpretation of concept names in $\mathcal{I}$. Therefore, we mostly define the SLO $\mathcal{F}^+$ for *frames* $\mathcal{F}$, interpretations in which $A^{\mathcal{F}} = \emptyset$ for all concept names A. One can apply algebraic notation to interpretations in a straightforward way. For example, we say that a SLO equation or quasi-equation is *valid* in $\mathcal{I}$ if it is valid in the algebra $\mathcal{I}^+$. It is known that natural contraints on the interpretation of roles can be captured by the validity of SLO equations. For example, if r, s are role names, then r is included in s in an interpretation $\mathcal{I}$ iff the equation $\Diamond_r x \leq \Diamond_s x$ is valid in $\mathcal{I}$. Formally, call a set P of first-order sentences using role names as binary predicates a *role constraint*. Then we say that a role constraint P *corresponds* to a set Ax of SLO equations if any interpretation $\mathcal{I}$ satisfies P iff $\mathcal{I}^+ \models \alpha$, for all $\alpha \in \mathsf{Ax}$. The table below gives a sample set of role constraints and the corresponding SLO equations. These correspondences are well known from correspondence theory in modal logic and are, in particular, instances of the correspondence part of Sahlqvist's Theorem [4, 6]. We refer the reader to [29, 17] for more examples.

Role constraint	Equation
$\mathsf{func}(r)$	$\Diamond_r x \wedge \Diamond_r y \leq \Diamond_r(x \wedge y)$
$r \sqsubseteq s^-$	$x \wedge \Diamond_r y \leq \Diamond_r(y \wedge \Diamond_s x)$
$r_1 \circ \cdots \circ r_n \sqsubseteq r$	$\Diamond_{r_1} \cdots \Diamond_{r_n} x \leq \Diamond_r x$

The following example illustrates how we use these correspondences to determine approximations.

Example 6. Consider a $\mathcal{ELF}_\bot$ ontology $\mathcal{O}_E$ containing as its only role assertion $\mathsf{func}(r)$. Thus, $\mathcal{O}_E = \mathcal{O} \cup \{\mathsf{func}(r)\}$, for a set $\mathcal{O}$ of CIs. Set $\Sigma = \mathsf{sig}(\mathcal{O}_E)$. To approximate $\mathcal{O}_E$ by a $\mathcal{EL}_\bot$ ontology, we take the equation $\alpha = (\Diamond_r x \wedge \Diamond_r y \leq \Diamond_r(x \wedge y))$ corresponding to $\mathsf{func}(r)$ and replace in $\mathcal{O}_E$ the functionality assertion $\mathsf{func}(r)$ by the set of all CIs obtained from α^C by substituting A_x and A_y by arbitrary $\mathcal{EL}_\bot(\Sigma)$ concepts. Thus, we form the set α^Σ of all CIs $\exists r.C \sqcap \exists r.D \sqsubseteq \exists r.(C \sqcap D)$, where C, D are $\mathcal{EL}_\bot(\Sigma)$ concepts and claim that $\mathcal{O}_L = \mathcal{O} \cup \alpha^\Sigma$ is an $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_E$.

It is easy to see that correpondence of $\mathsf{func}(r)$ to α entails that $\mathcal{O}_E \models \mathcal{O}_L$. The converse direction ($\mathcal{O}_E \models C \sqsubseteq D$ implies $\mathcal{O}_L \models C \sqsubseteq D$ for all $\mathcal{EL}_\bot(\Sigma)$ CIs $C \sqsubseteq D$), however, does *not* follow from correspondence and requires significantly more work - which we discuss next.

We develop a necessary and sufficient condition for when equations Ax corresponding to a role constraint P provide a $\mathcal{EL}_\bot$ approximation of ontologies of the form $\mathcal{O} \cup P$, where $\mathcal{O}$ is a set of $\mathcal{EL}_\bot$ CIs.

A *homomorphism* h between SLOs $\mathfrak{A}_1$ and $\mathfrak{A}_2$ is a mapping h from the domain A_1 of $\mathfrak{A}_1$ to the domain A_2 of $\mathfrak{A}_2$ preserving all operations, for example $h(\Diamond_r^{\mathfrak{A}_1} a) = \Diamond_r^{\mathfrak{A}_2} h(a)$ for all $a \in A_1$ and $r \in \mathcal{R}$. An *embedding* is an injective homomorphism. A set Ax of SLO equations is *complex* if for every SLO $\mathfrak{A}$ validating Ax there exists a frame $\mathcal{F}$ validating Ax such that $\mathfrak{A}$ can be embedded into $\mathcal{F}^+$. Thus, if Ax is complex, then every SLO $\mathfrak{A}$ validating Ax can be regarded as a system of sets (aka complexes) over a frame validating Ax. Call Ax *quasi-equation complete* if a quasi-equation is valid in all SLOs validating Ax just in case it is valid in all SLOs of the form $\mathcal{I}^+$ validating Ax. It can be proved that a set Ax of SLO equations is complex iff it is quasi-equation complete [17]. Theorem 7 can be proved by showing that the *empty set* of SLO equations is complex. In other words, the *empty* role constraint P corresponds to the *empty* complex set Ax of equations. For an equation $\alpha = (\sigma \leq \tau)$ we denote by α^Σ the set of all CIs obtained from α^C by uniformly substituting every A_x in α^C by any $\mathcal{EL}_\bot(\Sigma)$ concept D. Let Ax^Σ denote the union of all α^Σ, $\alpha \in \mathsf{Ax}$. We are now in the position to formulate the announced criterion for approximations. Observe that the approximations only depend on the role constraint of the ontology and not on its CIs.

Theorem 8. *[Approximation/Axiomatization] Let P be a role constraint and* Ax *a set of SLO equations corresponding to P. Then the following conditions are equivalent:*

1. $\mathcal{O} \cup \mathsf{Ax}^\Sigma$ *is a $\mathcal{EL}_\bot$ approximation of $\mathcal{O} \cup P$, for all $\mathcal{EL}_\bot$ ontologies $\mathcal{O}$ with* $\Sigma = \mathsf{sig}(\mathcal{O})$*;*
2. Ax *is complex;*
3. Ax *is quasi-equation complete.*

Proof. We first show that (2) implies (1). Assume Ax is complex and let $\mathcal{O}$ be an $\mathcal{EL}_\bot$ ontology and with $\Sigma = \mathsf{sig}(\mathcal{O} \cup P)$. We have to show that for all $\mathcal{EL}_\bot(\Sigma)$ concepts C, D: $\mathcal{O} \cup P \models C \sqsubseteq D$ iff $\mathcal{O} \cup \mathsf{Ax}^\Sigma \models C \sqsubseteq D$. The direction ($\Leftarrow$) follows from the observation that every interpretation satisfying P validates Ax and, therefore, satisfies all CIs in Ax^Σ. Conversely, assume that $\mathcal{O} \cup \mathsf{Ax}^\Sigma \not\models C \sqsubseteq D$. Take an interpretation $\mathcal{I}$ satisfying $\mathcal{O} \cup \mathsf{Ax}^\Sigma$ such that $\mathcal{I} \not\models C \sqsubseteq D$. Let $\mathcal{R}$ be the set of role names in Σ. Define the SLO $\mathfrak{A} = (A, \wedge^{\mathfrak{A}}, \bot^{\mathfrak{A}}, \top^{\mathfrak{A}}, (\Diamond_r^{\mathfrak{A}} \mid r \in \mathcal{R}))$

as the restriction of the SLO $\mathcal{I}^+$ to $\{C^{\mathcal{I}} \mid \mathsf{sig}(C) \subseteq \Sigma\}$. In more detail,

$$\begin{aligned} A &= \{C^{\mathcal{I}} \mid \mathsf{sig}(C) \subseteq \Sigma\} \\ X \wedge^{\mathfrak{A}} Y &= X \cap Y \\ \top^{\mathfrak{A}} &= \top^{\mathcal{I}} \\ \bot^{\mathfrak{A}} &= \emptyset \\ \Diamond_r^{\mathfrak{A}} X &= \{d \in \Delta^{\mathcal{I}} \mid \exists d' \in X \ (d, d') \in r^{\mathcal{I}}\} \end{aligned}$$

Then $\mathfrak{A}$ validates Ax: to see this let v be a valuation in $\mathfrak{A}$. By definition, for every variables x there exists a $\mathcal{EL}_\bot(\Sigma)$ concept C_x with $v(x) = C_x^{\mathcal{I}}$. Let $\sigma \leq \tau \in \mathsf{Ax}$. Obtain σ^s and τ^s from σ^C and τ^C by substituting every A_x by C_x. Then $\sigma^s \sqsubseteq \tau^s \in \mathsf{Ax}^\Sigma$. Thus $\mathcal{I} \models \sigma^s \sqsubseteq \tau^s$ and so $\mathfrak{A} \models_v \sigma \leq \tau$, as required.

As Ax is complex, there exists a frame $\mathcal{G}$ validating Ax such that there is an embedding h from $\mathfrak{A}$ into $\mathcal{G}^+$. Extend $\mathcal{G}$ to an interpretation $\mathcal{J}$ by setting $A^{\mathcal{J}} = h(A^{\mathcal{I}})$ for every concept name A. Then $\mathcal{J}$ is a model of $\mathcal{O}$ validating Ax and refuting $C \sqsubseteq D$. Thus, $\mathcal{J}$ is a model of $\mathcal{O}$ and P and refuting $C \sqsubseteq D$.

We now show that (1) implies (3). Assume Ax is not quasi-equation complete. Take $\mathfrak{A}$ validating Ax and a quasi-equation $\alpha_1 \wedge \cdots \wedge \alpha_n \to \alpha$ such that

(a) $\mathfrak{A} \not\models \alpha_1 \wedge \cdots \wedge \alpha_n \to \alpha$;
(b) $\mathcal{I} \models \alpha_1 \wedge \cdots \wedge \alpha_n \to \alpha$ for all $\mathcal{I}$ validating Ax.

We may assume that all variables and $\Diamond_r$ used in α are used in some α_i. Let v be a valuation in $\mathfrak{A}$ such that $\alpha_1, \ldots, \alpha_n$ are true in $\mathfrak{A}$ under v and α is refuted in $\mathfrak{A}$ under v. Take a frame $\mathcal{F}$ such that $\mathfrak{A}$ is embedded into $\mathcal{F}^+$ via an injective homomorphism h. Define a model $\mathcal{J}$ by expanding $\mathcal{F}$ by setting $A_x^{\mathcal{J}} = h(v(x))$, for all variables x in $\alpha_1, \ldots, \alpha_n, \alpha$. Let $\mathcal{O} = \{\alpha_1^T, \ldots, \alpha_n^T\}$ and $\Sigma = \mathsf{sig}(\mathcal{O} \cup P)$. Then $\mathcal{I}$ is a model of Ax^Σ (since $\mathfrak{A}$ validates Ax) and $\mathcal{I} \not\models \alpha^T$. It follows that $\mathcal{O} \cup \mathsf{Ax}^\Sigma \not\models \alpha^T$. However, by Point (b), $\mathcal{O} \cup P \models \alpha^T$.

The equivalence (2) $\Leftrightarrow$ (3) is proved in [17]. ❏

We now exhibit role constraints given by role assertions and inclusions and corresponding axioms that are complex. An application of Theorem 8 then provides the desired approximations. Let P be a set of role assertions and inclusions. Recall that we call P *inverse closed* if for every role name r in P there is a role name $\hat{r}$ such that $r \sqsubseteq \hat{r}^-, \hat{r} \sqsubseteq r^- \in P$ and there are no additional occurrences of inverse roles in P. Recall that we assume that P is *safe* in the sense that for any assertion $\rho_1 \circ \cdots \circ \rho_n \sqsubseteq \rho \in P$ with $n \geq 2$ neither $P \models \mathsf{func}(\rho)$ nor $P \models \mathsf{func}(\rho^-)$ holds.

Theorem 9. *Let P be an inverse closed and safe set of role assertions and RIs and let Ax contain the following equations, for all role names r, s in P:*

1. $\Diamond_r x \leq \Diamond_s x$ *if* $P \models r \sqsubseteq s$;
2. $\Diamond_r x \wedge \Diamond_r y \leq \Diamond_r (x \wedge y)$ *if* $\mathsf{func}(r) \in P$;
3. $x \wedge \Diamond_r y \leq \Diamond_r (y \wedge \Diamond_{\hat{r}} x)$;
4. $\Diamond_r \Diamond_{\hat{r}} x \leq x$ *if* $\mathsf{func}(\hat{r}) \in P$;

5. $\Diamond_{r_1} \cdots \Diamond_{r_n} x \leq \Diamond_r x$ *and* $\Diamond_{\hat{r}_n} \cdots \Diamond_{\hat{r}_1} x \leq \Diamond_{\hat{r}} x$ *if* $r_1 \circ \cdots \circ r_n \sqsubseteq r \in P$.

Then P *corresponds to* Ax *and* Ax *is complex.*

Proof. Correpondence is straightforward, so we focus on proving that Ax is complex. We start by introducing some equations implied by Ax. We set $P \models^* \mathsf{func}(r)$ if there exists s such that $P \models r \sqsubseteq s$ and $P \models \mathsf{func}(s)$.

(A) $\mathsf{Ax} \models \Diamond_r x \wedge \Diamond_s \top \leq \Diamond_s x$ if $P \models^* \mathsf{func}(r)$ and $P \models s \sqsubseteq r$. To see this, assume $P \models r \sqsubseteq u$ and $\mathsf{func}(u) \in P$. Then

$$\mathsf{Ax} \models \Diamond_r x \wedge \Diamond_s \top \leq \Diamond_u x \wedge \Diamond_s \top$$

by the equations under Point (1.). We have by the equations under Point (3.)

$$\mathsf{Ax} \models \Diamond_u x \wedge \Diamond_s \top \leq \Diamond_s (\top \wedge \Diamond_{\hat{s}} \Diamond_u x)$$

Thus,

$$\mathsf{Ax} \models \Diamond_u x \wedge \Diamond_s \top \leq \Diamond_s \Diamond_{\hat{s}} \Diamond_u x$$

By the equations under Point (1.)

$$\mathsf{Ax} \models \Diamond_u x \wedge \Diamond_s \top \leq \Diamond_s \Diamond_{\hat{u}} \Diamond_u x$$

By the equations under Point (4.)

$$\mathsf{Ax} \models \Diamond_u x \wedge \Diamond_s \top \leq \Diamond_s x$$

We obtain

$$\mathsf{Ax} \models \Diamond_r x \wedge \Diamond_s \top \leq \Diamond_s x,$$

as required.

(B) $\mathsf{Ax} \models \Diamond_r x \wedge \Diamond_r y \leq \Diamond_r (x \wedge y)$ if $P \models^* \mathsf{func}(r)$. To see this, assume $P \models r \sqsubseteq u$ and $\mathsf{func}(u) \in P$. Then, by the equations under Point (1.)

$$\mathsf{Ax} \models \Diamond_r x \wedge \Diamond_r y \leq \Diamond_u x \wedge \Diamond_u y$$

By the equations under Point (2.)

$$\mathsf{Ax} \models \Diamond_u x \wedge \Diamond_u y \leq \Diamond_u (x \wedge y)$$

Thus, by (A),

$$\mathsf{Ax} \models \Diamond_r x \wedge \Diamond_r y \leq \Diamond_r (x \wedge y)$$

(C) $\mathsf{Ax} \models \Diamond_r \Diamond_{\hat{r}} x \leq x$ if $P \models^* \mathsf{func}(\hat{r})$. This follows from the equations under Point (1.) and (4.).

We introduce some notation for the proof. Let $\mathfrak{A}$ be a SLO. A *filter* F in $\mathfrak{A}$ is a subset of A such that $\top^{\mathfrak{A}} \in F$, $\perp^{\mathfrak{A}} \notin F$, and for all $a, b \in A$: $b \in F$ if $a \in F$ and $a \leq^{\mathcal{A}} b$, and $a \wedge b \in F$ if $a, b \in F$.

Using $\mathfrak{A}$, we define a frame $\mathcal{F}$ such that there is an embedding from $\mathfrak{A}$ to $\mathcal{F}^+$. Let $\Delta^{\mathfrak{F}}$ be the set of all filters in $\mathfrak{A}$. For any role name r, the definition of $r^{\mathcal{F}}$ depends on whether r and/or r^* are functional:

– if $P \models^* \mathsf{func}(r)$ and $P \models^* \mathsf{func}(\hat{r})$, then $(F_1, F_2) \in r^{\mathcal{F}}$ if (i) $\Diamond_r a \in F_1$ iff $a \in F_2$ and (ii) $\Diamond_{\hat{r}} a \in F_2$ iff $a \in F_1$. We interpret $\hat{r}$ as the inverse of r.
– if $P \models^* \mathsf{func}(r)$ and $P \not\models^* \mathsf{func}(\hat{r})$, then $(F_1, F_2) \in r^{\mathcal{F}}$ if (i) $\Diamond_r a \in F_1$ iff $a \in F_2$ and (ii) $\Diamond_{\hat{r}} a \in F_2$ if $a \in F_1$. We interpret $\hat{r}$ as the inverse of r.
– if $P \not\models^* \mathsf{func}(r)$ and $P \not\models^* \mathsf{func}(\hat{r})$, then $(F_1, F_2) \in r^{\mathcal{F}}$ iff (i) $\Diamond_r a \in F_1$ if $a \in F_2$ and (ii) $\Diamond_{\hat{r}} a \in F_2$ if $a \in F_1$. We interpret $\hat{r}$ as the inverse of r.

This finishes the definition of $\mathcal{F}$. We first show that $\mathcal{F}$ satisfies P.

– Assume $r \sqsubseteq s \in P$. We have to check that $r \sqsubseteq s$ is satisfied in $\mathcal{F}$. If neither $P \models^* \mathsf{func}(s)$ nor $P \models^* \mathsf{func}(\hat{s})$, then $r^{\mathcal{F}} \subseteq s^{\mathcal{F}}$ follows directly from the definition and the equations under Point (1).
  Now assume that $P \models^* \mathsf{func}(s)$. Then $P \models^* \mathsf{func}(r)$. Assume $(F_1, F_2) \in r^{\mathcal{F}}$. We have to show $(F_1, F_2) \in s^{\mathcal{F}}$. Thus, we first have to show that $\Diamond_s a \in F_1$ iff $a \in F_2$. If $a \in F_2$, then $\Diamond_r a \in F_1$. Then $\Diamond_s a \in F_1$ by the equations under Point (1), as required. If $\Diamond_s a \in F_1$, then $\Diamond_r a \in F_1$ since $\Diamond_r \top \in F_1$ (by the equations under (A)). Then $a \in F_2$, as required. Next we make a case distinction: if $P \not\models^* \mathsf{func}(\hat{s})$, then we have to show that $\Diamond_{\hat{r}} a \in F_2$ if $a \in F_1$. But this follows from $P \models \hat{r} \sqsubseteq \hat{s}$ and the equations under Point (1). If $P \models^* \mathsf{func}(\hat{s})$, then we have to show that $\Diamond_{\hat{s}} a \in F_2$ iff $a \in F_1$. This can be proved again using the equations under Point (1) and (A).
  The case $P \models^* \mathsf{func}(\hat{s})$ is considered in the same way.
– Assume $\mathsf{func}(r) \in P$. Then functionality of $r^{\mathcal{F}}$ follows directly from the definition.
– Assume $r_1 \circ \cdots \circ r_n \sqsubseteq r \in P$. Let $(F_1, F_2) \in (r_1 \circ \cdots \circ r_n)^{\mathcal{I}}$. As P is safe, we have to show that if $\Diamond_r a \in F_1$, then $a \in F_2$ and if $\Diamond_{\hat{r}} a \in F_2$, then $a \in F_1$. Both can be proved in a straightforward way using the equations under Point (5).

It remains to construct an embedding h from $\mathfrak{A}$ into $\mathcal{F}^+$. Define h by setting

$$h(a) := \{F \in \Delta^{\mathcal{F}} \mid a \in F\},$$

for all $a \in A$. It is straightforward to show that h is an injective mapping with

– $h(\top^{\mathfrak{A}}) = \Delta^{\mathcal{F}}$;
– $h(\bot^{\mathfrak{A}}) = \emptyset$;
– $h(a \wedge^{\mathfrak{A}} b) = h(a) \cap h(b)$.

It thus remains to prove that

$$h(\Diamond_r^{\mathfrak{A}} a) = \Diamond_r^{\mathcal{F}^+} h(a)$$

for all role names r. We first assume that $P \models^* \mathsf{func}(r)$ and $P \models^* \mathsf{func}(\hat{r})$. Assume a filter F is given. Suppose $\Diamond_r a_0 \in F$. We have to show the existence of a filter F' with $a_0 \in F'$ such that $(F, F') \in r^{\mathcal{F}}$. Consider

$$X = \{a \mid \Diamond_r a \in F\} \cup \{\Diamond_{\hat{r}} b \mid b \in F\}$$

and

$$Y = \{a \mid \Diamond_r a \notin F\} \cup \{\Diamond_{\hat{r}} b \mid b \notin F\}$$

It suffices to show the existence of a filter F' containing X with an empty intersection with Y. To this end it suffices to prove that there is no finite conjunction c of members of X such that $c \leq e$ for some $e \in Y$. Assume an arbitrary such c is given. It takes the form

$$c = a_1 \wedge \cdots \wedge a_n \wedge \Diamond_{\hat{r}} b_1 \wedge \cdots \Diamond_{\hat{r}} b_m$$

with $\Diamond_r a_1, \ldots, \Diamond_r a_n \in F$ and $b_1, \ldots, b_m \in F$. Then, by the axioms under Point (B), we have $\Diamond_r(a_1 \wedge \cdots \wedge a_n) \in F$. We also have $b_1 \wedge \cdots \wedge b_m \in F$. Thus we may assume that

$$c = a \wedge \Diamond_{\hat{r}} b$$

for some a with $\Diamond_r a \in F$ and $b \in F$. For a proof by contraction first assume that

$$a \wedge \Diamond_{\hat{r}} b \leq a'$$

for some a' with $\Diamond_r a' \notin F$. Then

$$\Diamond_r(a \wedge \Diamond_{\hat{r}} b) \leq \Diamond_r a'$$

by monotonicity of $\Diamond_r$. But by the equations under Point (3),

$$b \wedge \Diamond_r a \leq \Diamond_r(a \wedge \Diamond_{\hat{r}} b)$$

Then from $b \wedge \Diamond_r a \in F$ (since $b, \Diamond_r a \in F$) and $b \wedge \Diamond_r a \leq \Diamond_r a'$ we obtain $\Diamond_r a' \in F$ and we have derived a contradiction.

Now assume

$$a \wedge \Diamond_{\hat{r}} b \leq \Diamond_{\hat{r}} b'$$

for some $b' \notin F$. Then, by the equations under Point (3) and Point (C) and monotonicity of $\Diamond_r$,

$$b \wedge \Diamond_r a \leq \Diamond_r(a \wedge \Diamond_{\hat{r}} b) \leq \Diamond_r \Diamond_{\hat{r}} b' \leq b'$$

which contradicts the assumptions that $b, \Diamond_r a \in F$ and $b' \notin F$.

Now assume that $P \models^* \mathsf{func}(r)$ and $P \not\models^* \mathsf{func}(\hat{r})$. Assume a filter F is given. Suppose first $\Diamond_r a_0 \in F$. We have to show the existence of a filter F' with $a_0 \in F'$ such that $(F, F') \in r^{\mathcal{F}}$. Consider

$$X = \{a \mid \Diamond_r a \in F\} \cup \{\Diamond_{\hat{r}} b \mid b \in F\}$$

and

$$Y = \{a \mid \Diamond_r a \notin F\}$$

It suffices to show the existence of a filter F' containing X with an empty intersection with Y. To this end it suffices to prove that there is no finite conjunction

c of members of X such that $c \leq e$ for some $e \in Y$. Assume an arbitrary such c is given. As shown above, we may assume that

$$c = a \wedge \Diamond_{\hat{r}} b$$

for some a with $\Diamond_r a \in F$ and $b \in F$. Now one can prove that

$$a \wedge \Diamond_{\hat{r}} b \leq a'$$

for some a' with $\Diamond_r a' \notin F$ leads to a contradiction in exactly the same way as above.

Suppose now that $\Diamond_{\hat{r}} a_0 \in F$. We have to show the existence of a filter F' with $a_0 \in F'$ such that $(F', F) \in r^{\mathcal{F}}$. Consider

$$X = \{\Diamond_r a \mid a \in F\} \cup \{a_0\}$$

and

$$Y = \{\Diamond_r a \mid a \notin F\} \cup \{b \mid \Diamond_{\hat{r}} b \notin F\}$$

It suffices to show the existence of a filter F' containing X with an empty intersection with Y. To this end it suffices to prove that there is no finite conjunction c of members of X such that $c \leq e$ for some $e \in Y$. We may assume that $c = a_0 \wedge \Diamond_r a$ for some $a \in F$. Assume that

$$a_0 \wedge \Diamond_r a \leq \Diamond_r a'$$

for some $a' \notin F$. Then by the equations under Point (3) and (C)

$$a \wedge \Diamond_{r^*} a_0 \leq \Diamond_{\hat{r}}(a_0 \wedge \Diamond_r a) \leq \Diamond_{\hat{r}} \Diamond_r a' \leq a'$$

which contradicts the assumption that $a, \Diamond_{\hat{r}} a \in F$.

Assume that

$$a_0 \wedge \Diamond_r a \leq b$$

for some b with $\Diamond_{\hat{r}} b \notin F$. Then by the equations under Point (3)

$$a \wedge \Diamond_{\hat{r}} a_0 \leq \Diamond_{\hat{r}}(a_0 \wedge \Diamond_r a) \leq \Diamond_{\hat{r}} b$$

which again contradicts the assumption that $a, \Diamond_{\hat{r}} a \in F$.

The remaining case in which $P \not\models^* \mathsf{func}(r)$ and $P \not\models^* \mathsf{func}(\hat{r})$ is similar and omitted. ❏

Theorems 9 and 8 provide us with an $\mathcal{EL}_\perp$ approximation $\mathcal{O}_L$ of any inverse closed $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\perp$ ontology $\mathcal{O}_E$.

Theorem 10. *Let $\mathcal{O}_E$ be an inverse closed $\mathcal{ELR}^{\mathcal{I}}\mathcal{F}_\perp$ ontology and $\Sigma = \mathsf{sig}(\mathcal{O}_E)$. Define $\mathcal{O}_L$ as the $\mathcal{EL}_\perp$ ontology containing for all $\mathcal{EL}(\Sigma)$ concepts C, D and role names $r, s \in \Sigma$:*

1. *all CIs in $\mathcal{O}_E$;*

2. $\exists r.C \sqsubseteq \exists s.C$ *if* $\mathcal{O}_E \models r \sqsubseteq s$*;*
3. $\exists r_1.\cdots\exists r_n.C \sqsubseteq \exists r.C, \exists\hat{r}_n.\cdots\exists\hat{r}_1.C \sqsubseteq \exists\hat{r}.C$*, if* $r_1 \circ \cdots \circ r_n \sqsubseteq r \in \mathcal{O}_E$ *with* $n \geq 2$*;*
4. $C \sqcap \exists r.D \sqsubseteq \exists r.(D \sqcap \exists\hat{r}.C)$*;*
5. $\exists r.C \sqcap \exists r.D \sqsubseteq \exists r.(C \sqcap D)$ *if* $\mathsf{func}(r) \in \mathcal{O}_E$*;*
6. $\exists r.\exists\hat{r}.C \sqsubseteq C$ *if* $\mathsf{func}(\hat{r}) \in \mathcal{O}_E$*.*

Then $\mathcal{O}_L$ *is an* $\mathcal{EL}_\bot$ *approximation of* $\mathcal{O}_E$*.*

Theorem 1 is an immediate consequence of Theorem 10: if $\mathcal{O}_E$ is the $\mathcal{EL}_\bot$ approximation given by Theorem 10 and $\mathcal{O}'_E$ is the $\mathcal{ELR}_\bot$ ontology given in Theorem 1, then $\mathcal{O}'_E \models \mathcal{O}_E$ since

$$\{r \sqsubseteq s\} \models \exists r.C \sqsubseteq \exists s.C$$

and

$$\{r_1 \circ \cdots \circ r_n \sqsubseteq r\} \models \exists r_1.\cdots\exists r_n.C \sqsubseteq \exists r.C$$

and

$$\{r_1 \circ \cdots \circ r_n \sqsubseteq r\} \models \exists\hat{r}_n.\cdots\exists\hat{r}_1.C \sqsubseteq \exists\hat{r}.C.$$

The following examples show that correspondence between role assertions and SLO axioms alone does not imply that the axioms are complex and, therefore, cannot be used to obtain $\mathcal{EL}_\bot$ approximations. They also show that approximations are not compositional.

Example 7. Recall that the role assertion $P_0 = \{r \sqsubseteq s^-\}$ corresponds to the axiom $\alpha_0 = (x \wedge \Diamond_r y \leq \Diamond_r(y \wedge \Diamond_s x))$. Using the technique of the proof of Theorem 9, it is straightforward to show that α_0 is complex. Thus, Theorem 8 provides a $\mathcal{EL}_\bot$ approximation of any ontology $\mathcal{O} \cup P_0$ with $\mathcal{O}$ a set of $\mathcal{EL}_\bot$ inclusions. This is not the case if one admits two such role assertions. To show this, consider

$$P_1 = \{r_1 \sqsubseteq r_2^-, r_2 \sqsubseteq r_3^-\}$$

and let

$$\mathcal{O}_1 = \{A \sqsubseteq \exists r_1.B\} \cup P_1$$

Then P_1 corresponds to

$$\mathsf{Ax}_1 = \{x \wedge \Diamond_{r_1} y \leq \Diamond_{r_1}(y \wedge \Diamond_{r_2} x), x \wedge \Diamond_{r_2} y \leq \Diamond_{r_2}(y \wedge \Diamond_{r_3} x)\}$$

However, $\mathcal{O}'_1 = \{A \sqsubseteq \exists r_1.B\} \cup \mathsf{Ax}_1^{\{r_1,r_2,r_3,A,B\}}$ is not a $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_1$. To prove this, observe that

$$\mathcal{O}_1 \models A \sqsubseteq r_3.B.$$

We show that $\mathcal{O}' \not\models A \sqsubseteq \exists r_3.B$. Define an interpretation $\mathcal{I}$ by setting

- $\Delta^\mathcal{I} = \{a, b, c, d\}$;
- $r_1^\mathcal{I} = \{(a, b), (c, d)\}$;
- $r_2^\mathcal{I} = \{(b, c), (d, c)\}$;

- $r_3^{\mathcal{I}} = \{(c, d)\}$;
- $A^{\mathcal{I}} = \{a, c\}$, $B^{\mathcal{I}} = \{b, d\}$.

Then $\mathcal{I}$ is a model of $\mathcal{O}_1'$ but $a \in A^{\mathcal{I}} \setminus (\exists r_3.B)^{\mathcal{I}}$.

Observe that one obtains a $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_1$ by adding the axiom $\Diamond_{r_1} x \leq \Diamond_{r_3} x$ corresponding to $r_1 \sqsubseteq r_3$ to Ax_1.

The next example also refutes compositionality of approximations. In this case for combinations of RIs with functionality assertions.

Example 8. Let

$$P_2 = \{r \sqsubseteq s^-, \mathsf{func}(s)\}$$

Both role assertions in P_2 correspond to complex axioms, namely

$$\alpha_2 = (x \wedge \Diamond_r y \leq \Diamond_r (y \wedge \Diamond_s x))$$

and

$$\alpha_3 = (\Diamond_s x \wedge \Diamond_s y \leq \Diamond_s (x \wedge y)),$$

respectively. $\mathsf{Ax}_2 = \{\alpha_2, \alpha_3\}$ is not complex, however, and

$$\mathcal{O}_2' = \{\top \sqsubseteq \exists r.\top, \top \sqsubseteq \exists s.A\} \cup \mathsf{Ax}_2^{\{r,s,A\}}$$

is not a $\mathcal{EL}_\bot$ approximation of

$$\mathcal{O}_2 = \{\top \sqsubseteq \exists r.\top, \top \sqsubseteq \exists s.A\} \cup P_2$$

To show this, first observe that $\mathcal{O}_2 \models \top \sqsubseteq A$. To prove this, let $\mathcal{I}$ be a model of $\mathcal{O}_2$ and assume $(d, d') \in r^{\mathcal{I}}$. Then $(d', d) \in s^{\mathcal{I}}$ and from functionality of s and $\top \sqsubseteq \exists s.A$ we obtain $d \in A^{\mathcal{I}}$. On the other hand, $\mathcal{O}_2' \not\models \top \sqsubseteq A$: consider the interpretation $\mathcal{I}$ with domain $\Delta^{\mathcal{I}} = \{d, d'\}$, $r^{\mathcal{I}} = s^{\mathcal{I}} = \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ and $A^{\mathcal{I}} = \{d\}$. Then $\mathcal{I}$ is a model of $\mathcal{O}_2'$ but $d' \notin A^{\mathcal{I}}$.

One can define a $\mathcal{EL}_\bot$ approximation of $\mathcal{O}_2$ by adding the axiom $\Diamond_r \Diamond_s x \leq x$ (corresponding to the equations under Point (4)) to Ax_2.

Finally, we give an example illustrating why the equations under Point (4) are needed.

Example 9. Let

$$P_3 = \{t \sqsubseteq r, \mathsf{func}(r)\}$$

Both role assertions in P_2 correspond to complex axioms, namely

$$\alpha_4 = (\Diamond_t x \leq \Diamond_r x)$$

and

$$\alpha_5 = (\Diamond_r x \wedge \Diamond_r y \leq \Diamond_r (x \wedge y)),$$

respectively. $\mathsf{Ax}_3 = \{\alpha_4, \alpha_5\}$ is not complex, however, and

$$\mathcal{O}_3' = \{\exists r.A \sqcap \exists t.\top \sqsubseteq \exists t.A\} \cup \mathsf{Ax}_3^{\{t,r,A\}}$$

is not a $\mathcal{EL}_\bot$ approximation of

$$\mathcal{O}_3 = \{\exists r.A \sqcap \exists t.\top \sqsubseteq \exists t.A\} \cup P_3$$

To show this, first observe that $\mathcal{O}_3 \models \exists r.A \sqcap \exists t.\top \sqsubseteq \exists t.A$. On the other hand, $\mathcal{O}'_3 \not\models \exists r.A \sqcap \exists t.\top \sqsubseteq \exists t.A$. To see this, consider the interpretation $\mathcal{I}$ defined by setting

- $\Delta^\mathcal{I} = \{a, b, c\}$;
- $r^\mathcal{I} = \{(a, b)\}$, $t^\mathcal{I} = \{(a, c)\}$, $A^\mathcal{I} = \{b\}$.

Then $\mathcal{I}$ is a model of $\mathcal{O}'_3$ but $a \notin (\exists t.A)^\mathcal{I}$.

C Details for Section 4

C.1 The Chase

We start with introducing ABoxes, which the chase procedure uses as a data structure. Let $\mathsf{N_I}$ be a countably infinite set of *individual names* disjoint from $\mathsf{N_C}$ and $\mathsf{N_R}$. An *ABox* is a finite set of *concept assertions* $A(a)$, and *role assertions* $r(a, b)$ where $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and $a, b \in \mathsf{N_I}$. We use $\mathtt{Ind}(\mathcal{A})$ to denote the set of individual names that occur in the ABox $\mathcal{A}$. An interpretation $\mathcal{I}$ *satisfies* a concept assertion $A(a)$ if $a \in A^\mathcal{I}$ and a role assertion $r(a, b)$ if $(a, b) \in r^\mathcal{I}$. Note that we adopt the standard names assumption here, which implies the unique name assumption. An interpretation if a *model* of an ABox if it satisfies all assertions in it. For an ontology $\mathcal{O}$, ABox $\mathcal{A}$, $a \in \mathtt{Ind}(\mathcal{A})$, and $\mathcal{ELI}$ concept C, we write $\mathcal{A}, \mathcal{O} \models C(a)$ if $a \in C^\mathcal{I}$ for every model $\mathcal{I}$ of $\mathcal{A}$ and $\mathcal{O}$. Moreover, $\mathcal{A}$ is *consistent* with $\mathcal{O}$ if $\mathcal{A}$ and $\mathcal{O}$ have common model. We write $\mathcal{A} \models C(a)$ if $a \in C^\mathcal{I}$ where $\mathcal{I}$ is $\mathcal{A}$ viewed as an interpretation in the obvious way. An ABox $\mathcal{A}$ is *ditree-shaped* if the directed graph $G_\mathcal{A} = (\mathtt{Ind}(\mathcal{A}), \{(a, b) \mid r(a, b) \in \mathcal{A}\}$ is a tree; note that multi-edges are admitted.

Let $\mathcal{O}$ be an $\mathcal{ELFH}^\mathcal{I}_\bot$ ontology. We can assume w.l.o.g. that the $\bot$ concept occurs only in assertions of the form $C \sqsubseteq \bot$ with C an $\mathcal{EL}$ concept. Starting from a ditree-shaped ABox $\mathcal{A}$, the chase exhaustively applies the following rules:

- R1 If $\mathcal{A} \models C(a)$ and $C \sqsubseteq D \in \mathcal{O}$ with $D \neq \bot$, then add $D(a)$ to $\mathcal{A}$;
- R2 If $r(a, b) \in \mathcal{A}$ and $\mathcal{O} \models r \sqsubseteq s$ with r, s role names, then add $s(a, b)$ to $\mathcal{A}$
- R3 If $r(a, b) \in \mathcal{A}$, $\mathcal{O} \models r \sqsubseteq s^-$, and $\mathcal{A} \models C(a)$ with $\exists s.C$ a subconcept of $\mathcal{O}$ or an $\mathcal{EL}(\Sigma)$ concept of depth bounded by $\ell - 1$, then add $\exists s.C(b)$ to $\mathcal{A}$;[5]
- R4 If $r_1(a, b_1), r_2(a, b_2) \in \mathcal{A}$, $\mathcal{A} \models C_1(b_1)$, $\mathcal{A} \models C_2(b_2)$, C_1, C_2 $\mathcal{EL}(\Sigma)$ concepts of depth bounded by $\ell - 1$, $\mathcal{O} \models r_1 \sqsubseteq s$, $\mathcal{O} \models r_2 \sqsubseteq s$, and $\mathtt{func}(s) \in \mathcal{O}$, then add $\exists r_1.(C_1 \sqcap C_2)(a)$ to $\mathcal{A}$.

Note that the rule R3 is parameterized by a depth bound ℓ that is assumed to be identical to the depth bound ℓ used in the construction of $\mathcal{O}_L$. It can be verified that when the chase is started on a ditree-shaped ABox, then all ABoxes

[5] If $\ell = 0$, then there are no concepts of the latter form.

produced are ditree-shaped; in particular, 'backwards edges' as enforced by role inclusions of the form $r \sqsubseteq s^-$ are not made explicit, but only treated implicitly.

The chase applies the above rules exhaustively in a fair way. We assume that rules are *not* applied when its post-condition is already satisfied. For example, R3 is not applied when $\mathcal{A} \models \exists s.C(b)$. Using database theory parlance, one could say that our chase is *not oblivious*. This has the (undesired) consequence that the result of the chase, obtained in the limit by exhaustive and fair rule application, is not unique as it depends on the order in which rules are applies. However, all possible results are homomorphically equivalent and for the constructions in this paper, it does not matter which of the many possible results we use. For simplicity, we thus pretend that the outcome of the chase is unique and denote it with $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. The (desired) consequence of not being oblivious is that the (infinite) ABox $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ has finite outdegree.

The following lemma implies that the chase is (sound and) complete regarding consequences formulated in terms of $\mathcal{EL}$ concepts of depth bounded by ℓ. It is, however, incomplete regarding deeper $\mathcal{EL}$ concepts and regarding consequences formulated in $\mathcal{ELI}$.

Lemma 3. *Let $\mathcal{O}$ be an $\mathcal{ELFH}^{\mathcal{I}}_{\bot}$ ontology and $\mathcal{A}$ a ditree-shaped ABox with root a_0. Then*

1. *$\mathcal{A}$ is inconsistent with $\mathcal{O}$ iff there are $C \sqsubseteq \bot \in \mathcal{O}$ and $a \in \mathtt{Ind}(\mathcal{A})$ with $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(a)$, and*
2. *if $\mathcal{A}$ is consistent with $\mathcal{O}$, then $\mathcal{A}, \mathcal{O} \models C_0(a_0)$ iff $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C_0(a_0)$ for all $\mathcal{EL}$ concepts of depth at most ℓ.*

Proof. We consider both points simultaneously. The "if" directions are straightforward. In fact, it suffices to show that whenever an ABox $\mathcal{A}'$ is obtained from a ditree-shaped ABox $\mathcal{A}$ by application of one of the rules, then every model of $\mathcal{A}$ and $\mathcal{O}$ is also a model of $\mathcal{A}'$ and $\mathcal{O}$. This is straightforward using a case distinction according to which rule is applied and easy semantic arguments.

For the (contrapositive of the) "only if" directions, assume that there are no $C \sqsubseteq \bot \in \mathcal{O}$ and $a \in \mathtt{Ind}(\mathcal{A})$ such that $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(a)$, respectively that $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a_0)$. We show how to construct a model $\mathcal{J}$ of $\mathcal{A}$ and $\mathcal{O}$ such that for all $a \in \mathtt{Ind}(\mathcal{A})$ and $\mathcal{EL}$ concepts C, $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a)$ implies $a \notin C^{\mathcal{J}}$. This implies that $\mathcal{A}$ is consistent with $\mathcal{O}$ (since $\bot$ occurs in $\mathcal{O}$ only in the form $C \sqsubseteq \bot$), respectively that $\mathcal{A}, \mathcal{O} \not\models C(a_0)$.

Let $\sim$ be the smallest equivalence relation on the individuals in $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ such that whenever $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ contains $r(a, b_1)$ and $r(a, b_2)$ with $\mathtt{func}(r)$, then $b_1 \sim b_2$. Clearly, for any equivalence class of $\sim$, there is an individual a such that all individuals of the class are successors of a in the ditree $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. We call a the *predecessor* of the class. An individual b is *maximal* if for every b' with $b \sim b'$ and every $\mathcal{EL}$ concept C of depth at most $\ell - 1$ with $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(b')$, we have $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(b)$.

Claim 1. Every equivalence class of $\sim$ contains a maximal individual.

It is clear that the outdegree of the ditree $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ is finite and thus each equivalence class $\{b_1, \dots, b_k\}$ of $\sim$ is finite since all individuals in it have a common predecessor. Assume that the class does not contain a maximal individual. Then there must be b_{i_1}, b_{i_2} in the class and $\mathcal{EL}$ concepts C_1, C_2 of depth at most $\ell - 1$ such that $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C_1(b_{i_1})$, $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C_2(b_{i_2})$, and there is no b_{i_3} in the class with $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C_1(b_{i_3})$ and $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C_2(b_{i_3})$. But this situation is impossible since $\mathsf{R4}$ was applied exhaustively.

Let $\mathcal{A}_1$ be obtained by closing $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ as follows:

($\dagger_1$) whenever $r(a, b) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ and $\mathcal{O} \models r \sqsubseteq s^-$, then add $s(b, a)$.

Note that while $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$, $\mathcal{A}_1$, and $\mathcal{A}_2$ only have downwards edges, ($\dagger_3$) only adds upwards edges. Moreover, due to the assumed syntactic restriction ($\heartsuit$), when ($\dagger_3$) adds $s(b, a)$, then $\mathtt{func}(s) \notin \mathcal{O}$ and $\mathtt{func}(s^-) \notin \mathcal{O}$. We say that ($\dagger_3$) does *not add functional edges* respectively *does not add inverse functional edges.*

Next, let $\mathcal{A}_2$ be obtained from $\mathcal{A}_1$ as follows:

($\dagger_2$) whenever b_1 and b_2 are successors of a with $b_1 \sim b_2$ and $\rho(a, b_2) \in \mathcal{A}_1$, then add $\rho(a, b_1)$.

Note that ($\dagger_2$) may add both upwards and downwards edges, but it does not add functional or inverse functional upwards edges. In fact, let r_{ab_2} be the *primary* role name between a and b_2, that is, $t(a, b_2) \in \mathcal{A}_0$ implies $\mathcal{O} \models r_{ab_2} \sqsubseteq t$. Such a role name must exist by definition of the chase: r_{ab_i} is the role name from the first edge that the chase has introduced between a and b_2 and all remaining edges were added later by $\mathsf{R2}$. Now observe that $r_{ab_2} \sqsubseteq \rho$ and thus if ρ is an inverse role then by ($\heartsuit$) it can neither be functional nor inverse functional.

Finally, let $\mathcal{A}_3$ be obtained from $\mathcal{A}_2$ as follows:

($\dagger_3$) for every individual a and every $\sim$-equivalence class $\{b_1, \dots, b_k\}$ of which a is the predecessor: choose a maximum individual b_i and remove all edges $r(a, b_j)$ and subtrees rooted at b_j, $j \neq i$.

For brevity, let $\mathcal{A}_0 = \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. We prove the following central claim:

Claim 2. For every $\mathcal{EL}$ concept C that is a subconcept of $\mathcal{O}$ or of depth bounded by ℓ, every $a \in \Delta^{\mathcal{I}}$, and $i \in \{0, 1, 2\}$, $\mathcal{A}_i \models C(a)$ iff $\mathcal{A}_{i+1} \models C(a)$.

We distinguish the cases $i \in \{0, 1, 2\}$. In all cases, the proof is by induction on the structure of C and the only interesting case is that C is of the form $\exists s.D$.

Case $i = 0$. Since the “only if” direction is clear, we concentrate on “if”. Assume that $\mathcal{A}_1 \models \exists s.D(a)$ and let $s(a, b) \in \mathcal{A}_1$ with $b \in \mathcal{A}_1 \models D(b)$. The induction hypothesis yields $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models D(b)$. If $s(a, b) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$, then clearly $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models \exists s.D(a)$. If this is not the case, then $s(a, b)$ was added by ($\dagger_1$). Then b is a predecessor of a and there is $r(b, a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ such that $\mathcal{O} \models r \sqsubseteq s^-$. Since $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models D(b)$, $\mathsf{R3}$ was applied resulting in $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models \exists s.D(a)$.

Case $i = 1$. Since the “only if” direction is clear, we concentrate on “if”. First assume that ρ is a role name s. Thus assume that $\mathcal{A}_2 \models \exists s.D(a)$ and let

$s(a,b) \in \mathcal{A}_2$ with $\mathcal{A}_2 \models D(b)$. The induction hypothesis and Case $i = 0$ yield $\mathcal{A}_0 \models D(b)$. The interesting case is that $s(a,b)$ was added to $\mathcal{A}_2$ by ($\dagger_2$). Then b is a successor of an individual a and a has another successor b_2 such that $s(a,b_2) \in \mathcal{A}_1$ and $b \sim b_2$. Since ($\dagger_1$) adds only upwards edges, $s(a,b_2) \in \mathcal{A}_0$. Since $b \sim b_2$, there are individuals $c_1, \dots, c_k$ and role names $r_1, \dots, r_{k-1}$ such that

- $c_1 = b$ and $c_k = b_2$
- $r_i(a,c_i), r_i(a,c_{i+1}) \in \mathcal{A}_0$ for $1 \leq i < k$
- $\mathtt{func}(r_1), \dots, \mathtt{func}(r_{k-1}) \in \mathcal{O}$.

For $1 \leq i \leq k$, let r_{ac_i} be the *primary* role name between a and c_i. We must have

- $\mathcal{O} \models r_{ac_i} \sqsubseteq r_i(a,c_i)$ for $1 \leq i < k$
- $\mathcal{O} \models r_{ac_i} \sqsubseteq r_i(a,c_{i-1})$ for $1 < i \leq k$.

We can thus apply R4 $k-1$ times to obtain an individual b' such that $r_{ac_k}(a,b') \in \mathcal{A}_0$ and $\mathcal{A}_0 \models D(b')$; note in this context that D is of depth at most $\ell - 1$. We already know that $\mathcal{O} \models r_{ac_k} \sqsubseteq s$ and thus R2 yields $s(a,b') \in \mathcal{A}_0$ which implies $\mathcal{A}_0 \models \exists s.D(a)$ and thus $\mathcal{A}_1 \models \exists s.D(a)$ as required. The case where ρ is an inverse role s^- is similar. In fact, the statement "$s(a,b_2) \in \mathcal{A}_0$" is then replaced with "$t(a,b_2) \in \mathcal{A}_0$ for some role name t with $t \sqsubseteq s^-$". We can use the same argument as above and add at the end that ($\dagger_1$) has added $s(b',a)$.

Case $i = 2$. Here, the "if" direction is clear and we concentrate on "only if". Thus assume that $\mathcal{A}_2 \models \exists s.D(a)$ and let $s(a,b) \in \mathcal{A}_2$ with $\mathcal{A}_2 \models D(b)$. The interesting case is when $s(a,b)$ and the subtree below b was removed by ($\dagger_3$). By Cases $i = 0$ and $i = 1$, $\mathcal{A}_0 \models D(b)$. By Claim 1, there is a maximal b' from the equivalence class of b and thus $\mathcal{A}_0 \models D(b')$ implying $\mathcal{A}_2 \models D(b')$. Because of ($\dagger_2$), $s(a,b') \in \mathcal{A}_2$ and thus $\mathcal{A}_2 \models \exists s.D(a)$ as required. This finishes the proof of Claim 2.

Let $\mathcal{I}$ be $\mathcal{A}_3$ viewed as an interpretation. We first argue that $\mathcal{I}$ satisfies all role inclusions in $\mathcal{O}$. Thus let $r \sqsubseteq \rho \in \mathcal{I}$ and $(a,b) \in r^{\mathcal{I}}$. Then $r(a,b) \in \mathcal{A}_3$. First assume that $r(a,b) \in \mathcal{A}_0$. Then $\rho(a,b) \in \mathcal{A}_0$ by R2 and ($\dagger_1$) and thus $\rho(a,b) \in \mathcal{A}_3$. The case that $r(a,b)$ was added by ($\dagger_1$) also relies on R2 and ($\dagger_1$), and the semantics. Now assume that $r(a,b)$ was added by ($\dagger_2$). There are two cases. Either b is a successor of a and a has another successor b_2 such that $b \sim b_2$ and $r(a,b_2) \in \mathcal{A}_1$. But then $\rho(a,b_2) \in \mathcal{A}_1$ and thus ($\dagger_2$) adds also $s(a,b)$. Or a is a successor of b and b has another successor a_2 such that $a \sim a_2$ and $r(a,b_2) \in \mathcal{A}_1$. Again, ($\dagger_2$) adds also $s(a,b)$.

We next show that $\mathcal{I}$ satisfies all functionality assertions in $\mathcal{O}$. To see this, assume that $(a,b_1), (a,b_2) \in \rho^{\mathcal{I}}$ and $\mathtt{func}(\rho) \in \mathcal{O}$ where ρ is a role name or the inverse thereof. If ρ is an inverse role, we must have $b_1 = b_2$ as required: since $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ is ditree-shaped, for every individual a there is at most one b with $r(b,a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ for every role name r. Since ($\dagger_1$) and ($\dagger_2$) do not add inverse functional edges, the same is true for $\mathcal{A}_3$ when $\mathtt{func}(r^-) \in \mathcal{O}$. Now assume that ρ is a role name r. Since ($\dagger_1$) and ($\dagger_2$) do not add functional upwards edges, both edges $r(a,b_1), r(a,b_2)$ must also be in $\mathcal{A}_0$ and must thus be downwards edges. But then $(\dagger_3)$ ensures that $b_1 = b_2$.

It follows from Claim 2 that $\mathcal{I}$ satisfies all concept inclusions $C \sqsubseteq D \in \mathcal{O}$ with $D \neq \bot$. In fact, let $a \in C^{\mathcal{I}}$. Claim 2 yields $\texttt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(a)$, rule $\mathsf{R1}$ gives $\texttt{chase}_{\mathcal{O}}(\mathcal{A}) \models D(a)$ and applying Claim 2 once more gives $a \in D^{\mathcal{I}}$.

We now finish the proofs of the "only if" directions of Points 1 and 2 of Lemma 3. For Point 1, by assumption we have $\texttt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a)$ for all $a \in \texttt{Ind}(\mathcal{A})$ and $C \sqsubseteq \bot \in \mathcal{O}$, and thus Claim 2implies that all concept inclusions $C \sqsubseteq \bot \in \mathcal{O}$ are satisfied by $\mathcal{I}$. Thus, $\mathcal{I}$ is a model of $\mathcal{O}$, which shows that $\mathcal{A}$ $\mathcal{A}$ is consistent with $\mathcal{O}$, finishing the argument.

For Point 2, by assumption we have that $\mathcal{A}$ is consistent with $\mathcal{O}$ and that $\texttt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a_0)$. From the former and the already established "if" direction of Point 1, we get $\texttt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a)$ for all $a \in \texttt{Ind}(\mathcal{A})$ and $C \sqsubseteq \bot \in \mathcal{O}$. Thus, $\mathcal{I}$ is again a model of $\mathcal{O}$ and Claim 2 yields $a_0 \notin C_0^{\mathcal{I}}$, thus $\mathcal{A}, \mathcal{O} \not\models C_0(a_0)$ as required. ❑

C.2 Completeness

We now prove the completeness part of Theorem 2, starting with some preliminaries. For an ABox $\mathcal{A}$ and an interpretation $\mathcal{I}$, a function $h : \texttt{Ind}(\mathcal{A}) \to \Delta^{\mathcal{I}}$ is a *homomorphism from $\mathcal{A}$ to $\mathcal{I}$* if $h(a) \in A^{\mathcal{I}}$ for every $A(a) \in \mathcal{A}$ and $(h(a), h(b)) \in r^{\mathcal{I}}$ for every $r(a,b) \in \mathcal{A}$. For two ABoxes $\mathcal{A}_1$ and $\mathcal{A}_2$, a function $h : \texttt{Ind}(\mathcal{A}_1) \to \texttt{Ind}(\mathcal{A}_2)$ is a *homomorphism from $\mathcal{A}_1$ to $\mathcal{A}_2$* if $A(h(a)) \in \mathcal{A}_2$ for every $A(a) \in \mathcal{A}_1$ and $r(h(a), h(b)) \in \mathcal{A}_2$ for every $r(a,b) \in \mathcal{A}_1$. We recall that every $\mathcal{EL}$ concept C can be viewed as a ditree-shaped ABox $\mathcal{A}_C$. By convention, we assume that the root individual in such an ABox is a_0. For example, the $\mathcal{EL}$ concept $A \sqcap \exists r.B \sqcap \exists s.\top$ can be viewed as the ABox $\{A(a_0), r(a_0,b_1), B(b_1), s(a_0,b_2)\}$. The following is widely known and straightforward to establish.

Lemma 4. *Let C be an $\mathcal{EL}$ concept, $\mathcal{I}$ an interpretation, and $d \in \Delta^{\mathcal{I}}$. Then $d \in C^{\mathcal{I}}$ iff there is a homomorphism h from $\mathcal{A}_C$ to $\mathcal{I}$ with $h(a_0) = d$.*

The next lemma is the central step in the completeness proof. It says that the chase introduced above can in a sense be simulated by the statements in the ontology $\mathcal{O}_L$.

Lemma 5. *Let C be an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ decorated with subconcepts of $\mathcal{O}_E$ at leaves, $\mathcal{A}_0, \mathcal{A}_1, \ldots$ the sequence of ABoxes constructed by $\texttt{chase}_{\mathcal{O}_E}(\mathcal{A}_C)$, $\mathcal{I}$ a model of $\mathcal{O}_L$, and $d \in C^{\mathcal{I}}$. Then for every $i \geq 0$, there is a homomorphism h_i from $\mathcal{A}_i$ to $\mathcal{I}$ with $h_i(a_0) = d$.*

Proof. The proof is by an induction on the number of applications of the chase rule $\mathsf{R3}$ used to compute the sequence $\mathcal{A}_0, \ldots, \mathcal{A}_i$. For the induction start, assume that $\mathsf{R3}$ was not applied at all. We show that for all $j \leq i$, there is a homomorphism h_j from $\mathcal{A}_j$ to $\mathcal{I}$ with $h_j(a_0) = d$. For $j = 0$, it suffices to apply

Lemma 4. Now assume that h_j has already been constructed, $j < i$. We show how to find h_{j+1}, making a case distinction according to the rule that is applied in order to obtain $\mathcal{A}_{j+1}$ from $\mathcal{A}_j$.

R1. Then $\mathcal{A}_j \models C(a)$, $C \sqsubseteq D \in \mathcal{O}_E$, and $\mathcal{A}_{j+1}$ is obtained from $\mathcal{A}_j$ by adding $D(a)$. We must have $h_j(a) \in C^{\mathcal{I}}$. Since $C \sqsubseteq D$ is a CI in $\mathcal{O}_L$ and $\mathcal{I}$ is a model of $\mathcal{O}_L$, $a \in D^{\mathcal{I}}$. Thus h_j can be extended to a homomorphism from $\mathcal{A}_{j+1}$ to $\mathcal{I}$ in a straightforward way.

R2. Then $r(a,b) \in \mathcal{A}_j$, $\mathcal{O}_E \models r \sqsubseteq s$, and $\mathcal{A}_{j+1}$ is obtained from $\mathcal{A}_j$ by adding $s(a,b)$. We have $r \sqsubseteq s \in \mathcal{O}_L$ and thus $h_{j+1} = h_j$ is a homomorphism from $\mathcal{A}_{j+1}$ to $\mathcal{I}$.

R4. There are $r_1(a,b_1), r_2(a,b_2) \in \mathcal{A}_j$ such that $\mathcal{A}_j \models C_1(b_1)$, $\mathcal{A}_j \models C_2(b_2)$, C_1, C_2 $\mathcal{EL}(\Sigma)$ concepts of depth bounded by $\ell - 1$, $\mathcal{O}_E \models r_1 \sqsubseteq s$, $\mathcal{O}_E \models r_1 \sqsubseteq s$, $\texttt{func}(s) \in \mathcal{O}_E$, and $\mathcal{A}_{j+1}$ is obtained from $\mathcal{A}_j$ by adding $\exists r_1.C_1 \sqcap C_2(a)$. Clearly, $h_j(a) \in (\exists r_1.C_1 \sqcap \exists r_2.C_2)^{\mathcal{I}}$. By construction, $\mathcal{O}_L$ contains $\exists r_1.C_1 \sqcap \exists r_2.C_2 \sqsubseteq \exists r_1.(C_1 \sqcap C_2)$. Consequently, $h_j(a) \in \exists r_1.(C_1 \sqcap C_2)^{\mathcal{I}}$. We can thus extend h_j to the desired homomorphism h_{j+1} from $\mathcal{A}_{j+1}$ to $\mathcal{I}$ in a straightforward way.

Now for the induction step. Assume that there were $k > 0$ applications of R3 in the sequence $\mathcal{A}_0, \dots, \mathcal{A}_i$ and that the last such application was used to obtain $\mathcal{A}_{p+1}$ from $\mathcal{A}_p$, $p < i$. By induction hypothesis, we find a homomorphism h_p from $\mathcal{A}_p$ to $\mathcal{I}$ with $h_p(a_0) = d$. We argue that we also find such a homomorphism h_{p+1} from $\mathcal{A}_{p+1}$ to $\mathcal{I}$. We can then proceed as in the induction start to obtain the desired homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$.

Since R3 was applied, there are $r(a,b) \in \mathcal{A}_p$, a role name s, and an $\mathcal{EL}$ concept C' with $\exists s.C'$ a subconcept of $\mathcal{O}_E$ or of depth bounded by ℓ such that $\mathcal{A}_p \models C'(a)$, $\mathcal{O}_E \models r \sqsubseteq s^-$, and $\mathcal{A}_{p+1}$ is obtained from $\mathcal{A}_p$ by adding $\exists s.C'(b)$. Since $\mathcal{A}_p$ is ditree-shaped, a is the predecessor of b. By definition of the chase, there is a primary role name r_{ab} between a and b, as in the proof of Lemma 3: for every $t(a,b) \in \mathcal{A}_p$, we have $\mathcal{O}_E \models r_{ab} \sqsubseteq t$ and thus also $\mathcal{O}_L \models r_{ab} \sqsubseteq t$. It suffices to show that there is a $d_b \in \Delta^{\mathcal{I}}$ such that $(h_p(a), d_b) \in r_{ab}^{\mathcal{I}}$ and a homomorphism h_b from $\mathcal{A}_{p+1}|_b$ to $\mathcal{I}$ with $h_b(b) = d_b$, where $\mathcal{A}_{p+1}|_b$ is the restriction of $\mathcal{A}_{p+1}$ to the subtree rooted at b. In fact, it is then straightforward to combine h_p and h_b into the desired homomorphism h_{p+1}: set $h_{p+1}(c) = h_p(c)$ if c is not in $\mathcal{A}_{p+1}|_b$ and $h_{p+1}(c) = h_b(c)$ otherwise. Observe that $t(a,b) \in \mathcal{A}_{p+1}$ implies $(h(a), d_b) \in t^{\mathcal{I}}$ for all role names t since $(h_p(a), d_b) \in r_{ab}^{\mathcal{I}}$, $\mathcal{I}$ is a model of $\mathcal{O}_L$, and no new edges between a and b have been added in the construction of $\mathcal{A}_{p+1}$ from $\mathcal{A}_p$. In particular, $\mathcal{O}_L \models r_{ab} \sqsubseteq r$.

For brevity, let $\mathcal{A}_b = \mathcal{A}_p|_b$ and set $\texttt{Ind} = \{b\} \cup (\texttt{Ind}(\mathcal{A}_b) \cap \texttt{Ind}(\mathcal{A}_C))$. That is, $\texttt{Ind}$ contains only those individuals from $\mathcal{A}_b$ that were present already in the initial ABox $\mathcal{A}_C$, and if there is no such individual, then $\texttt{Ind} = \{b\}$. An individual $c \in \texttt{Ind}$ is a *fringe individual* if there is some $t(c,c') \in \mathcal{A}_p^b$ with $c' \notin \texttt{Ind}$. Further, let $\mathcal{A}_p^{b-}$ be the restriction of $\mathcal{A}_p^b$ to assertions that only use individuals from $\texttt{Ind}$ extended by adding $E(c)$ whenever E is a subconcept of $\mathcal{O}_E$ and c is a fringe individual such that $\mathcal{A}_p \models E(c)$, and let C_b be this ABox viewed as an $\mathcal{EL}$ concept.

We must have $h_p(b) \in C_b^{\mathcal{I}}$ and thus $h_p(a) \in (C' \sqcap \exists r_{ab}.C_b)^{\mathcal{I}}$. Since $\exists s.C'$ is a subconcept of $\mathcal{O}_E$ or of depth at most ℓ and C_b is an $\mathcal{EL}$ concept of depth at most ℓ' (by construction and because C is of depth bounded by ℓ; recall that $\ell' = \max\{\ell - 1, 0\}$) decorated with subconcepts of $\mathcal{O}_E$ at leaves, $C' \sqcap \exists r_{ab}.C_b \sqsubseteq \exists r_{ab}.(C_b \sqcap \exists s.C') \in \mathcal{O}_L$. Consequently, there is a $d_b \in (C_b \sqcap \exists s.C')^{\mathcal{I}}$ with $(d_a, d_b) \in r_{ab}^{\mathcal{I}}$. Let $\mathcal{B}$ be obtained from $\mathcal{A}_b^-$ by adding $\exists s.C'(b)$, as in the construction of $\mathcal{A}_{p+1}$. By Lemma 4, there is a homomorphism h_b from $\mathcal{B}$ to $\mathcal{I}$ with $h_b(b) = d_b$. It remains to extend h_b from $\mathcal{B}$ to $\mathcal{A}_{p+1}|_b$.

To this end, consider each fringe individual c. Let C_c be the $\mathcal{EL}$ concept that is the conjunction of all subconcepts E of $\mathcal{O}_E$ with $\mathcal{A}_b \models E(c)$. We can extract from the chase sequence $\mathcal{A}_0, \dots, \mathcal{A}_p$ a chase sequence that constructs $\mathcal{A}_p|_c$ starting from $\mathcal{A}_{C_c}$ and uses at most $k-1$ applications of special rules. From the induction hypothesis and since clearly $h_b(c) \in C_c^{\mathcal{I}}$, we thus obtain a homomorphism h_c from $\mathcal{A}_c$ to $\mathcal{I}$ with $h_c(c) = h_b(c)$. It is now straightforward to combine our initial h_b with all the homomorphisms h_c into the desired homomorphism h_b from $\mathcal{A}_{p+1}|_b$ to $\mathcal{I}$. ❏

We are now ready to prove completeness of the approximation constructed in Theorem 2. It is immediate by construction of $\mathcal{O}_L$ that $\mathcal{O}_E \models r \sqsubseteq s$ implies $\mathcal{O}_L \models r \sqsubseteq s$ for all role names $r, s \in \Sigma$. it thus remains to show the following.

Lemma 6. *$\mathcal{O}_E \models C \sqsubseteq D$ implies $\mathcal{O}_L \models C \sqsubseteq D$ for all $\mathcal{EL}(\Sigma)$ concepts C of depth at most ℓ and $\mathcal{EL}_\bot$ concepts D.*

Proof. By Point 2 of Lemma 1 and construction of $\mathcal{O}_L$, it suffices to consider the case $\ell < \omega$. Assume $\mathcal{O}_E \models C \sqsubseteq D$ with C, D as in Lemma 6 and let $\mathcal{I}$ be a model of $\mathcal{O}_L$ with $d \in C^{\mathcal{I}}$. We have to show that $d \in D^{\mathcal{I}}$. First assume that the ABox $\mathcal{A}_C$ is consistent with $\mathcal{O}_E$. By Point 2 of Lemma 3, $\mathtt{chase}_{\mathcal{O}_E}(\mathcal{A}_C) \models D(a_0)$. Let $\mathcal{A}_C = \mathcal{A}_0, \mathcal{A}_1, \dots$ be the sequence of ABoxes generated by the chase when started on $\mathcal{A}_C$. Then $\mathtt{chase}_{\mathcal{O}_E}(\mathcal{A}_C) \models D(a_0)$ implies that there is an $\mathcal{A}_i$ with $\mathcal{A}_i \models D(a_0)$. An analogue of Lemma 4 for homomorphisms into ABoxes thus yields a homomorphism h from $\mathcal{A}_D$ to $\mathcal{A}_i$ with $h(a_0) = a_0$ (defined in the expected way). By Lemma 5, there further is a homomorphism h' from $\mathcal{A}_i$ to $\mathcal{I}$ with $h(a_0) = d$. Composing these, we obtain a homomorphism from $\mathcal{A}_D$ to $\mathcal{I}$ that maps a_0 to d and applying Lemma 4 yields $d \in D^{\mathcal{I}}$, as required. Now assume that $\mathcal{A}_C$ is inconsistent with $\mathcal{O}_E$. Then by Point 1 of Lemma 3, there are $C' \sqsubseteq \bot \in \mathcal{O}_E$ and $a \in \mathtt{Ind}(\mathcal{A})$ with $\mathtt{chase}_{\mathcal{O}_E}(\mathcal{A}) \models C'(a)$. We can argue as above that there is a homomorphism from $\mathcal{A}_{C'}$ to $\mathcal{I}$, and thus ${C'}^{\mathcal{I}} \neq \emptyset$ in contradiction to the facts that $\mathcal{I}$ is a model of $\mathcal{O}_L$ and $C' \sqsubseteq \bot \in \mathcal{O}_L$. ❏

D Details for Section 5

In Section 5, we introduced two $\mathcal{ELI}_\bot$-to-$\mathcal{EL}_\bot$ approximations. In this section, we deliver further details for both cases.

D.1 The Chase (non-projective)

Let $\mathcal{O}$ be an $\mathcal{ELI}$ ontology and $\Sigma = \mathtt{sig}(\mathcal{O})$. We again assume that $\bot$ occurs only in CIs of the form $C \sqsubseteq \bot$. Starting from an ABox $\mathcal{A}$, the chase exhaustively applies the following rules:

R1 If $\mathcal{A} \models C(a)$ and $C \sqsubseteq D \in \mathcal{O}_E$ with $D \neq \bot$, then add $D(a)$ to $\mathcal{A}$.

The chase applies this rule exhaustively in a fair way. When the resulting sequence of ABoxes is $\mathcal{A} = \mathcal{A}_0, \mathcal{A}_1, \dots,$ we use $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ to denote $\bigcup_{i\geq 0} \mathcal{A}_i$. For simplicity, we assume here that the chase is oblivious in the sense that the rule applies even when the consequence $\mathcal{A} \models D(a)$ already holds. Consequently, $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ is uniquely defined. The following is easy to establish, details are omitted.

Lemma 7. *Let $\mathcal{O}$ be an $\mathcal{ELI}$ ontology and $\mathcal{A}$ an ABox. Then*

1. *$\mathcal{A}$ is inconsistent with $\mathcal{O}$ iff there are $C \sqsubseteq \bot \in \mathcal{O}$ and $a \in \mathtt{Ind}(\mathcal{A})$ with $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(a)$;*
2. *if $\mathcal{A}$ is consistent with $\mathcal{O}$, then $\mathcal{A}, \mathcal{O} \models C(a)$ iff $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \models C(a)$ for all $\mathcal{ELI}$ concepts C and $a \in \mathtt{Ind}(\mathcal{A})$.*

We are going to use the chase on ditree-shaped ABoxes. When started on such an ABox, all generated ABoxes are tree-shaped, that is, the undirected graph $G^u_{\mathcal{A}} = (\mathtt{Ind}(\mathcal{A}), \{\{a,b\} \mid r(a,b) \in \mathcal{A}\}$ is a tree (possibly with multi-edges); they are not guaranteed to be ditree-shaped.

D.2 Completeness (non-projective)

We prove the completeness part of Theorem 3. It is not difficult to prove that since $\mathcal{O}_E$ is formulated in $\mathcal{ELI}$, for any role inclusion $r \sqsubseteq s$, $\mathcal{O}_E \models r \sqsubseteq s$ implies $\mathcal{O}^{\mathcal{H}}_E \models r \sqsubseteq s$ where $\mathcal{O}^{\mathcal{H}}_E$ is the set of role inclusions from $\mathcal{O}_E$. Since $\mathcal{O}^{\mathcal{H}}_E \subseteq \mathcal{O}_L$, we have $\mathcal{O}_E \models r \sqsubseteq s$ iff $\mathcal{O}_L \models r \sqsubseteq s$ and it remains to deal with concept inclusions.

For a tree-shaped ABox $\mathcal{A}$, we use $\mathcal{A}^{\downarrow}$ to denote the restriction of $\mathcal{A}$ to assertions in which all individuals a are reachable from the root of $\mathcal{A}$ along a directed role path, that is, $\mathcal{A}$ contains assertions $r_0(a_0, a_1), \dots, r_{n-1}(a_{n-1}, a_n)$ where a_0 is the root of $\mathcal{A}$ and $a_n = a$.

Lemma 8. *Let C be an $\mathcal{EL}(\Sigma)$ concept, $\mathcal{A}_0, \mathcal{A}_1, \dots$ the sequence of ABoxes constructed by $\mathtt{chase}_{\mathcal{O}_E}(\mathcal{A}_C)$, $\mathcal{I}$ a model of $\mathcal{O}_L$, and $d \in C^{\mathcal{I}}$. Then for every $i \geq 0$, there is a homomorphism h_i from $\mathcal{A}^{\downarrow}_i$ to $\mathcal{I}$ with $h_i(a_0) = d$.*

Proof. The proof is by induction on i. The induction start is immediate by Lemma 4 and since $\mathcal{A}_0 = \mathcal{A}_C$. For the induction step, we make a case distinction according to the chase rule applied to obtain $\mathcal{A}_{i+1}$ from $\mathcal{A}_i$.

R1. Assume the rule was applied to individual $a \in \mathtt{Ind}(\mathcal{A}_i)$ and CI $E \sqsubseteq F \in \mathcal{O}_E$. Let $F^{\mathcal{EL}}$ be the result of replacing in F every subconcept $\exists r^-.G$ with $\top$. Moreover, let $\mathcal{A}^{\downarrow,-}_i$ be $\mathcal{A}^{\downarrow}_i$ after removal of all role edges that have been added by an application of rule R2, and likewise for $\mathcal{A}_{i+1}^{\downarrow,-}$. Further, let $C_i^{\downarrow,-}$ be $\mathcal{A}_i^{\downarrow,-}$ viewed as an $\mathcal{EL}$ concept, and likewise for $C_{i+1}^{\downarrow,-}$. Note that $\mathcal{A}_{i+1}^{\downarrow}$ is obtained from $\mathcal{A}_i^{\downarrow}$ by adding $F^{\mathcal{EL}}(a)$ and thus $C_{i+1}^{\downarrow,-}$ is a $\mathtt{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ decoration of $C_i^{\downarrow,-}$. From Point 2 of Lemma 7, we obtain $\mathcal{O}_E \models C \sqsubseteq C_{i+1}^{\downarrow,-}$; note that this (trivially) holds also when $\mathcal{A}_C$ is inconsistent with $\mathcal{O}_E$. Since $\emptyset \models C_i^{\downarrow,-} \sqsubseteq C$, this implies $\mathcal{O}_E \models C_i^{\downarrow,-} \sqsubseteq C_{i+1}^{\downarrow,-}$. As a consequence and since $C_{i+1}^{\downarrow,-}$ is a $\mathtt{cl}_{\mathcal{EL}}(\mathcal{O}_E)$ decoration of $C_i^{\downarrow,-}$, we must have $C_i^{\downarrow,-} \sqsubseteq C_{i+1}^{\downarrow,-} \in \mathcal{O}_L$. Thus, $d \in (C_{i+1}^{\downarrow,-})^{\mathcal{I}}$ and by Lemma 4 we find a homomorphism h_{i+1} from $\mathcal{A}_{i+1}^{\downarrow,-}$ to $\mathcal{I}$ with $h(a_0) = d$. Since $\mathcal{I}$ is a model of $\mathcal{O}$ and $\mathcal{O}_L$ contains the same role inclusions as $\mathcal{O}_E$, h_i must also be a homomorphism from $\mathcal{A}_{i+1}^{\downarrow}$ to $\mathcal{I}$, as required.

❏

Lemma 9. *Let C be an $\mathcal{EL}(\Sigma)$ concept and D an $\mathcal{EL}_\bot$ concept. Then $\mathcal{O}_E \models C \sqsubseteq D$ implies $\mathcal{O}_L \models C \sqsubseteq D$.*

Proof. Let $\mathcal{I}$ be a model of $\mathcal{O}_L$ with $d \in C^{\mathcal{I}}$. We have to show that $d \in D^{\mathcal{I}}$. Let $\mathcal{A}_0, \mathcal{A}_1, \dots$ be the sequences of ABoxes constructed by the chase started on $\mathcal{A}_C$. First assume that $\mathcal{A}_C$ is consistent with $\mathcal{O}_E$. Then from Point 2 of Lemma 7, we obtain $\mathtt{chase}_{\mathcal{O}_E}(\mathcal{A}_C) \models D(a_0)$. Thus, there is a k with $\mathcal{A}_k \models D(a_0)$. Since D is an $\mathcal{EL}$ concept, this implies $\mathcal{A}_k^{\downarrow} \models D(a_0)$. By Lemma 4, there is thus a homomorphism h_0 from $\mathcal{A}_D$ to $\mathcal{A}_k^{\downarrow}$ with $h_0(a_0) = a_0$. Lemma 8 yields a homomorphism h from $\mathcal{A}_k^{\downarrow}$ to $\mathcal{I}$ with $h(a_0) = d$. Composing h_0 with h and applying Lemma 4 yields $d \in D^{\mathcal{I}}$ as required. Now assume that $\mathcal{A}_C$ is inconsistent with $\mathcal{O}_E$. Then $\mathcal{O}_E \models C \sqsubseteq \bot$ and thus $C \sqsubseteq \bot \in \mathcal{O}_L$, in contradiction to $\mathcal{I}$ being a model of $\mathcal{O}_L$ with $C^{\mathcal{I}} \neq \emptyset$. ❏

D.3 The Chase (projective)

Let $\mathcal{O}$ be an $\mathcal{ELI}_\bot$ ontology in normal form. Starting from an ABox $\mathcal{A}$, the chase exhaustively applies the following rules, constructing in the limit an extended and potentially infinite ABox:

R1 If $A_1(a), \dots, A_n(a) \in \mathcal{A}$ and $\mathcal{O} \models A_1 \sqcap \dots \sqcap A_n \sqsubseteq B$, then add $B(a)$ to $\mathcal{A}$;
R2 If $r(a, b), A(b) \in \mathcal{A}$ and $\exists r.A \sqsubseteq B \in \mathcal{O}$, then add $B(a)$ to $\mathcal{A}$;
R3 If $r(b, a), A(b) \in \mathcal{A}$ and $\exists r^-.A \sqsubseteq B \in \mathcal{O}$, then add $B(a)$ to $\mathcal{A}$;
R4 If $A(a) \in \mathcal{A}$ and $A \sqsubseteq \exists r.B \in \mathcal{O}$, then add $r(a, b)$ and $B(b)$ to $\mathcal{A}$, with b fresh.

We again use $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ to denote the result of applying the chase of ABox $\mathcal{A}$, which is unique since rule application is oblivious. Note that this chase is the standard chase for $\mathcal{ELI}$ except that no new successors are introduced to witness existential restrictions on inverse roles. This is compensated by the semantic entailment in rule R1, which is in line with Point 1 in Theorem 4. We remark that, when applied to a ditree-shaped ABox, all ABoxes produced by the chase are ditree-shaped, possibly with multi-edges. We assume that $\bot$ occurs in $\mathcal{O}$ only in CIs of the form $C \sqsubseteq \bot$ with C an $\mathcal{EL}$ concept.

Lemma 10. *Let $\mathcal{O}$ be an $\mathcal{ELI}_\bot$ ontology and $\mathcal{A}$ a ditree-shaped ABox with root a_0. Then*

1. *$\mathcal{A}$ is inconsistent with $\mathcal{O}$ iff there are $C \sqsubseteq \bot \in \mathcal{O}$ and $a \in \mathtt{Ind}(\mathcal{A})$ with $\mathtt{chase}_\mathcal{O}(\mathcal{A}) \models C(a)$, and*
2. *if $\mathcal{A}$ is consistent with $\mathcal{O}$, then $\mathcal{A}, \mathcal{O} \models C_0(a_0)$ iff $\mathtt{chase}_\mathcal{O}(\mathcal{A}) \models C_0(a_0)$ for all $\mathcal{EL}$ concepts of depth at most ℓ.*

Proof. We prove both point simultaneously, starting with the "if" directions. Let $\mathcal{A}_0, \mathcal{A}_1, \dots$ be the sequence of ABoxes produced by the chase and let $\mathcal{I}$ be a model of $\mathcal{A}$ and $\mathcal{O}$. We show the following.

Claim For all $i \geq 0$, there is a homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$ with $h(a_0) = a_0$.

This establishes Point 1 because if $C \sqsubseteq \bot \in \mathcal{O}$ and $\mathtt{chase}_\mathcal{O}(\mathcal{A}) \models C(a)$, then there is a k such that $\mathcal{A}_k \models C(a)$. The existence of h_k implies that $C^\mathcal{I} \neq \emptyset$ (via Lemma 4), in contradiction to $\mathcal{I}$ being a model of $\mathcal{A}$ and $\mathcal{O}$. Thus, $\mathcal{A}$ is inconsistent with $\mathcal{O}$.

It also establishes Point 2. In fact, if $\mathtt{chase}_\mathcal{O}(\mathcal{A}) \models C(a_0)$, then there is a k such that $\mathcal{A}_k \models C(a_0)$ and h_k shows that $d \in C^\mathcal{I}$ as required.

It thus remains to prove the claim. The case $i = 0$ is trivial as the desired homomorphism is simply the identity on $\mathtt{Ind}(\mathcal{A})$. For the case $i > 0$, we make a case distinction according the the chase rule applied in order to obtain $\mathcal{A}_{i+1}$ from $\mathcal{A}_i$.

R1. If this rule was applied to obtain $\mathcal{A}_{i+1}$ from $\mathcal{A}_i$, then there are $A_1(a), \dots, A_n(a) \in \mathcal{A}_i$ such that $A_1 \sqcap \dots \sqcap A_n \sqsubseteq B \in \mathcal{O}_E$. $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ by adding $B(a)$. Because h_i is the homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$, we must have $h_i(a) \in (A_1 \sqcap \dots \sqcap A_n)^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{O}$, this yields $h_{i+1}(a) \in B^\mathcal{I}$. Consequently h_i is also a homomorphism from $\mathcal{A}_{i+1}$ to $\mathcal{I}$ and we can set $h_{i+1} = h_i$.

R2. Then there are $r(a, b), A(b) \in \mathcal{A}_i$ and $\exists r.A \sqsubseteq B \in \mathcal{O}$. $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ by adding $B(a)$. Because h_i is the homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$, we must have $h_i(b) \in A^\mathcal{I}$ and $(h_i(a), h_i(b)) \in r^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{O}$, $h_{i+1}(a) \in B^\mathcal{I}$. Consequently, h_i is also a homomorphism from $\mathcal{A}_{i+1}$ to $\mathcal{I}$ and we can set $h_{i+1} = h_i$.

R3. If this rule was applied to obtain $\mathcal{A}_{i+1}$ from $\mathcal{A}_i$, then there are $r(b, a), A(b) \in \mathcal{A}_i$ and $\exists r^-.A \sqsubseteq B \in \mathcal{O}$. $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ by adding $B(a)$. Because h_i is the homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$, we must have $h_i(b) \in A^\mathcal{I}$ and $(h_i(b), h_i(a)) \in r^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{O}$, this yields $h_{i+1}(a) \in B^\mathcal{I}$. Consequently, h_i is also a homomorphism from $\mathcal{A}_{i+1}$ to $\mathcal{I}$ and we can set $h_{i+1} = h_i$.

R4. Then there is an $A(a) \in \mathcal{A}_i$ such that $A \sqsubseteq \exists r.B \in \mathcal{O}$ and $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ by adding $B(b)$ and $r(a, b)$, b fresh. Because h_i is the homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$, we must have $h_i(a) \in A^\mathcal{I}$. Since $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{O}$, there is an $e \in B^\mathcal{I}$ such that $(h_i(a), e) \in r^\mathcal{I}$. Consequently $h_{i+1} = h_1 \cup \{b \mapsto e\}$ is a homomorphism from $\mathcal{A}_{i+1}$ to $\mathcal{I}$.

For the (contrapositive of the) "only if" directions, assume that there are no $C \sqsubseteq \bot \in \mathcal{O}$ and $a \in \mathtt{Ind}(\mathcal{A})$ such that $\mathtt{chase}_\mathcal{O}(\mathcal{A}) \models C(a)$, respectively that

$\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a_0)$. We show how to construct a model $\mathcal{J}$ of $\mathcal{A}$ and $\mathcal{O}$ such that for all $a \in \mathtt{Ind}(\mathcal{A})$ and $\mathcal{EL}$ concepts C, $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a)$ implies $a \notin C^{\mathcal{J}}$. This implies that $\mathcal{A}$ is consistent with $\mathcal{O}$ (since $\bot$ occurs in $\mathcal{O}$ only in the form $C \sqsubseteq \bot$), respectively that $\mathcal{A}, \mathcal{O} \not\models C(a_0)$.

For brevity, let $\mathtt{Ind}$ denote the set of individual names in the (potentially infinite) ABox $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. Further let $\mathcal{I}$ be $\mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ viewed as an interpretation. Because of rule $\mathsf{R1}$, for each $a \in \mathtt{Ind}$ we find a model $\mathcal{I}_a$ of $\mathcal{O}$ and a $d_a \in \Delta^{\mathcal{I}_a}$ such that for all concept names A, $a \in A^{\mathcal{I}}$ iff $d_a \in A^{\mathcal{I}_a}$. We can further assume that the domains of all interpretations $\mathcal{I}_a$ are mutually disjoint, and that they are also disjoint from the domain of $\mathcal{I}$. Let $\mathcal{J}$ be the interpretation obtained as follows:

1. take the disjoint union of $\mathcal{I}$ and all the $\mathcal{I}_a$;
2. for every $a \in \mathtt{Ind}$, every role name r, and every $(e, d_a) \in r^{\mathcal{I}_a}$, add (e, a) to $r^{\mathcal{J}}$.

It can be verified that, as required, $\mathtt{chase}_{\mathcal{O}}(\mathcal{A}) \not\models C(a)$ implies $a \notin C^{\mathcal{J}}$ for all $a \in \mathtt{Ind}(\mathcal{A})$ and $\mathcal{EL}$ concepts C. In fact, it suffices to observe that we have only added new incoming edges to elements from $\Delta^{\mathcal{I}}$ but no outgoing ones.

By definition, $\mathcal{J}$ is a model of $\mathcal{A}$. To show that it is also a model of $\mathcal{O}$, we make a case distinction on the types of CIs in $\mathcal{O}$:

$A_1 \sqcap \cdots \sqcap A_n \sqsubseteq B \in \mathcal{O}$. Let $a \in (A_1 \sqcap \cdots \sqcap A_n)^{\mathcal{J}}$. By construction of $\mathcal{J}$, $A_1(a), \ldots, A_n(a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. Thus, $\mathsf{R1}$ yields $B(a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ and $a \in B^{\mathcal{J}}$ by construction of $\mathcal{J}$.

$\exists r.A \sqsubseteq B \in \mathcal{O}$. Let $b \in A^{\mathcal{J}}$ and $(a, b) \in r^{\mathcal{J}}$. By construction of $\mathcal{J}$, $A(b)$ and $r(a, b) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. Thus, $\mathsf{R2}$ yields $B(a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ and $a \in B^{\mathcal{J}}$ by construction of $\mathcal{J}$.

$A \sqsubseteq \exists r.B \in \mathcal{O}$. Let $a \in A^{\mathcal{J}}$. By construction of $\mathcal{J}$, $A(a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. Thus, $\mathsf{R4}$ yields $r(a, b)$ and $B(b) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ and $b \in B^{\mathcal{J}}$ by construction of $\mathcal{J}$.

$\exists r^-.A \sqsubseteq B \in \mathcal{O}$. Let $b \in A^{\mathcal{J}}$ and $(b, a) \in r^{\mathcal{J}}$. By construction of $\mathcal{J}$, $A(b)$ and $r(b, a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$. Thus, $\mathsf{R3}$ yields $B(a) \in \mathtt{chase}_{\mathcal{O}}(\mathcal{A})$ and $a \in B^{\mathcal{J}}$ by construction of $\mathcal{J}$.

$A \sqsubseteq \exists r^-.B \in \mathcal{O}$. Let $a \in A^{\mathcal{J}}$. Because of $\mathsf{R1}$ we can find a model $\mathcal{I}_a$ with $(b, d_a) \in r^{\mathcal{I}_a}$. By point 2 of the construction of $\mathcal{J}$ we get $(b, a) \in r^{\mathcal{J}}, b \in B^{\mathcal{J}}$. ❏

D.4 Completeness (projective)

We prove the completeness part of Theorem 4, in analogy with the completeness proof for Theorem 2. Thus the following is crucial.

Lemma 11. *Let C be an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ, $\mathcal{A}_0, \mathcal{A}_1, \ldots$ the sequence of ABoxes constructed by $\mathtt{chase}_{\mathcal{O}_E}(\mathcal{A}_C)$, $\mathcal{I}$ a model of $\mathcal{O}_L$, and $d \in C^{\mathcal{I}}$. Then for every $i \geq 0$, there is a homomorphism h_i from $\mathcal{A}_i$ to $\mathcal{I}$ with $h_i(a_0) = d$.*

Proof. The proof is by an induction on the number of applications of rule R3 used to compute the sequence $\mathcal{A}_0, \dots, \mathcal{A}_i$. For the induction start, assume that R3 was not applied at all. We show that for all $j \leq i$, there is a homomorphism h_j from $\mathcal{A}_j$ to $\mathcal{I}$ with $h_j(a_0) = d$. For $j = 0$, it suffices to apply Lemma 4. Now assume that h_j has already been constructed, $j < i$. We show how to find h_{j+1}, making a case distinction according to the rule that is applied in order to obtain $\mathcal{A}_{j+1}$ from $\mathcal{A}_j$.

R1. Then there are $A_1(a), \dots, A_n(a) \in \mathcal{A}_j$ such that $A_1 \sqcap \dots \sqcap A_n \sqsubseteq B \in \mathcal{O}_E$ and $\mathcal{A}_{j+1}$ is obtained from $\mathcal{A}_j$ by adding $B(a)$. We must have $h_j(a) \in (A_1 \sqcap \dots \sqcap A_n)^{\mathcal{I}}$. By construction of $\mathcal{O}_L$, $A_1 \sqcap \dots \sqcap A_n \sqsubseteq B$ is a CI in $\mathcal{O}_L$. Since $\mathcal{I}$ is a model of $\mathcal{O}_L$, $a \in B^{\mathcal{I}}$. Thus h_j is also a homomorphism from $\mathcal{A}_{j+1}$ to $\mathcal{I}$.

R2. Then there are $r(a,b), A(b) \in \mathcal{A}_j$ such that $\exists r.A \sqsubseteq B \in \mathcal{O}$ and and $\mathcal{A}_{j+1}$ is obtained from $\mathcal{A}_j$ by adding $B(a)$. We must have $h_j(b) \in A^{\mathcal{I}}$ and $(h_j(a), h_j(b)) \in r^{\mathcal{I}}$. By construction of $\mathcal{O}_L$, $\exists r.A \sqsubseteq B$ is a CI in $\mathcal{O}_L$. Since $\mathcal{I}$ is a model of $\mathcal{O}_L$, $a \in B^{\mathcal{I}}$. Thus h_j is also a homomorphism from $\mathcal{A}_{j+1}$ to $\mathcal{I}$.

R4. Then there is an $A(a) \in \mathcal{A}_j$ such that $A \sqsubseteq \exists r.B \in \mathcal{O}$ and $\mathcal{A}_{j+1}$ is obtained from $\mathcal{A}_j$ by adding $B(b)$ and $r(a,b)$, b fresh. We must have $h_j(a) \in A^{\mathcal{I}}$. By construction of $\mathcal{O}_L$, $A \sqsubseteq \exists r.B$ is a CI in $\mathcal{O}_L$. Consequently and since $\mathcal{I}$ is a model of $\mathcal{O}_L$, there is an $e \in B^{\mathcal{I}}$ such that $(h_j(a), e) \in r^{\mathcal{I}}$. Let h_{j+1} be the extension of h_j obtained by setting $h_{j+1}(b) = e$. It can be verified that h_{j+1} is a homomorphism from $\mathcal{A}_{j+1}$ to $\mathcal{I}$.

Now for the induction step. Assume that there were $k > 0$ applications of R3 in the sequence $\mathcal{A}_0, \dots, \mathcal{A}_i$ and that the last such application was used to obtain $\mathcal{A}_{\ell+1}$ from $\mathcal{A}_\ell$, $\ell < i$. By induction hypothesis, we find a homomorphism h_ℓ from $\mathcal{A}_\ell$ to $\mathcal{I}$ with $h_\ell(a_0) = d$. We argue that we also find such a homomorphism $h_{\ell+1}$ from $\mathcal{A}_{\ell+1}$ to $\mathcal{I}$. We can then proceed as in the induction start to obtain the desired homomorphism from $\mathcal{A}_i$ to $\mathcal{I}$.

R3. There are $r(b,a), A(b) \in \mathcal{A}_\ell$ such that $\exists r^-.A \sqsubseteq B \in \mathcal{O}$ and $\mathcal{A}_{\ell+1}$ is obtained from $\mathcal{A}_\ell$ by adding $B(a)$. Since $\mathcal{A}_\ell$ is ditree-shaped, b is the predecessor of a in the tree $G_{\mathcal{A}_i}$. We must have $h_\ell(b) \in A^{\mathcal{I}}$ and $(h_\ell(b), h_\ell(a)) \in r^{\mathcal{I}}$. Let $\mathcal{A}_a$ be the ditree-shaped ABox in $\mathcal{A}_\ell$ rooted at a and set $\mathtt{Ind} = \{a\} \cup (\mathtt{Ind}(\mathcal{A}_a) \cap \mathtt{Ind}(\mathcal{A}_C))$. That is, $\mathtt{Ind}$ contains only those individuals in $\mathcal{A}_a$ that were present already in the initial ABox $\mathcal{A}_C$, and if there is no such individual, then $\mathtt{Ind} = \{a\}$. An individual $c \in \mathtt{Ind}$ is a *fringe individual* if there is some $r(c, c') \in \mathcal{A}_b$ with $c' \notin \mathtt{Ind}$. Further, let $\mathcal{A}_a^-$ be the restriction of $\mathcal{A}_a$ to assertions that only use individuals from $\mathtt{Ind}$ and let C_a be this ABox viewed as an $\mathcal{EL}$ concept.

We must have $h_\ell(a) \in C_a^{\mathcal{I}}$. By construction of $\mathcal{O}_L$ and since C_a is of depth at most $\ell - 1$ (because C is of depth bounded by ℓ), $A \sqcap \exists r.C_a \sqsubseteq \exists r.(C_a \sqcap B) \in \mathcal{O}_L$. Consequently, there is an $e \in (C_a \sqcap B)^{\mathcal{I}}$ with $(h_\ell(b), e) \in r^{\mathcal{I}}$. By Lemma 4, there is a homomorphism h_a from $\mathcal{A}_a^-$ to $\mathcal{I}$ with $h_a(a) = e$.

Now consider each fringe individual c. Let $\mathcal{A}_c$ be the ditree-shaped ABox in $\mathcal{A}_a$ rooted at c and let C_c be the $\mathcal{EL}$ concept that is the conjunction of all concept names A with $A(c) \in \mathcal{A}_a$. Since $\mathcal{O}_E$ is in normal form, we can extract from the chase sequence $\mathcal{A}_0, \dots, \mathcal{A}_\ell$ a chase sequence that constructs $\mathcal{A}_c$ starting from $\mathcal{A}_{C_c}$ and uses at most $k-1$ applications of R3. From the induction hypothesis and since clearly $h_a(c) \in C_c^{\mathcal{I}}$, we thus obtain a homomorphism h_c from $\mathcal{A}_c$ to $\mathcal{I}$ with $h_c(c) = h_a(c)$. We obtain the desired homomorphism $h_{\ell+1}$ from $\mathcal{A}_{\ell+1}$ to $\mathcal{I}$ by combining all these homomorphisms, that is,

$$h_{\ell+1}(c) = \begin{cases} h_\ell(c) & \text{if } c \notin \mathtt{Ind}(\mathcal{A}_a) \\ h_a(c) & \text{if } c \in \mathtt{Ind}(\mathcal{A}_a^-) \\ h_{c'}(c) & \text{if } c \in \mathtt{Ind}(\mathcal{A}_{c'}). \end{cases}$$

❑

The proof of the following lemma is now exactly identical to the proof of Lemma 6.

Lemma 12. *Let C be an $\mathcal{EL}(\Sigma)$ concept of depth bounded by ℓ and D an $\mathcal{EL}$ concept. Then $\mathcal{O}_E \models C \sqsubseteq D$ implies $\mathcal{O}_L \models C \sqsubseteq D$.*

E Details for Section 6

We require the following lemma from [23].

Lemma 13. *Asssume $\mathcal{O}$ is a $\mathcal{ELH}$ ontology and $\mathcal{O} \models C \sqsubseteq \exists r.D$. Then one of the following holds:*

1. *there exists a top-level conjunct $\exists s.C'$ of C such that $\mathcal{O} \models s \sqsubseteq r$ and $\mathcal{O} \models C' \sqsubseteq D$; or*
2. *there exists a subconcept M of $\mathcal{O}$ such that $\mathcal{O} \models C \sqsubseteq \exists r.M$ and $\mathcal{O} \models M \sqsubseteq D$.*

We use Lemma 13 to establish the results on non-finite and non-elementary approximations.

Theorem 5. *None of the ontologies*

$$\{\exists r^-.A \sqsubseteq B\}, \quad \{\mathsf{func}(r), A \sqsubseteq A\}, \quad \{r \sqsubseteq s^-, A \sqsubseteq A\}$$

has finite projective $\mathcal{ELH}$ approximations.

Proof. We start with $\mathcal{O}_E = \{\exists r^-.A \sqsubseteq B\}$. Let $\mathcal{O}_L$ be a projective $\mathcal{ELH}$ approximation of $\mathcal{O}_E$. For all $n \geq 0$, let $C_n = \exists r^n.\top$, where $\exists r^n$ denotes n-fold nesting of an existential restriction, and observe that

$$\mathcal{O}_E \models A \sqcap \exists r.C_n \sqsubseteq \exists r.(B \sqcap C_n).$$

To establish the desired result, it suffices to argue that for every $n \geq 0$, there is a subconcept M_n of $\mathcal{O}_L$ such that $\mathcal{O}_L \models M_n \sqsubseteq C_n$ and $\mathcal{O}_L \not\models M_n \sqsubseteq C_m$ for any $m > n$. First note that $\mathcal{O}_L \not\models C_n \sqsubseteq B$ for any $n \geq 0$ because the same is true for $\mathcal{O}_E$. By Lemma 13, to obtain $\mathcal{O}_L \models A \sqcap \exists r.C_n \sqsubseteq \exists r.(B \sqcap C_n)$, there must exist a subconcept M of $\mathcal{O}_L$ such that

– $\mathcal{O}_L \models A \sqcap \exists r.C_n \sqsubseteq \exists r.M$ and
– $\mathcal{O}_L \models M \sqsubseteq B \sqcap C_n$.

We aim to use M as M_n. Assume to the contrary of what remains to be shown that $\mathcal{O}_L \models M \sqsubseteq C_m$ for some $m > n$. Then $\mathcal{O}_L \models A \sqcap \exists r.C_n \sqsubseteq \exists r.(B \sqcap C_m)$, which contradicts the fact that this CI is not entailed by $\mathcal{O}_E$.

We now consider $\mathcal{O}_E = \{\mathsf{func}(r), A \sqsubseteq A\}$. Let $\mathcal{O}_L$ be a projective $\mathcal{ELH}$ approximation of $\mathcal{O}_E$. For all $n \geq 0$, let again $C_n = \exists r^n.\top$, and observe that

$$\mathcal{O}_E \models \exists r.C_n \sqcap \exists r.A \sqsubseteq \exists r.(C_n \sqcap A).$$

Using Lemma 13, we establish that for every $n \geq 0$, there is a subconcept M_n of $\mathcal{O}_L$ such that $\mathcal{O}_L \models M_n \sqsubseteq (C_n \sqcap A)$ and $\mathcal{O}_L \not\models M_n \sqsubseteq C_m$ for any $m > n$. First note that

– $\mathcal{O}_L \not\models C_n \sqsubseteq C_n \sqcap A$ for any $n \geq 0$ and
– $\mathcal{O}_L \not\models A \sqsubseteq C_n \sqcap A$

because the same is true for $\mathcal{O}_E$. To obtain $\mathcal{O}_L \models \exists r.C_n \sqcap \exists r.A \sqsubseteq \exists r.(C_n \sqcap A)$, there must exist a subconcept M of $\mathcal{O}_L$ such that $\mathcal{O}_L \models \exists r.C_n \sqcap \exists r.A \sqsubseteq \exists r.M$ and $\mathcal{O}_L \models M \sqsubseteq C_n \sqcap A$. We use M as M_n. Assume to the contrary of what remains to be shown that $\mathcal{O}_L \models M \sqsubseteq C_m$ for some $m > n$. Then $\mathcal{O}_L \models \exists r.C_n \sqcap \exists r.A \sqsubseteq \exists r.(C_m \sqcap A)$, which contradicts the fact that this CI is not entailed by $\mathcal{O}_E$.

We now consider $\mathcal{O}_E = \{r \sqsubseteq s^-, A \sqsubseteq A\}$. Let $\mathcal{O}_L$ be a projective $\mathcal{ELH}$ approximation of $\mathcal{O}_E$. For all $n \geq 0$, let again $C_n = \exists r^n.\top$, and observe that

$$\mathcal{O}_E \models A \sqcap \exists r.C_n \sqsubseteq \exists r.(C_n \sqcap \exists s.A).$$

Using Lemma 13, we establish that for every $n \geq 0$, there is a subconcept M_n of $\mathcal{O}_L$ such that $\mathcal{O}_L \models M_n \sqsubseteq C_n \sqcap \exists s.A$ and $\mathcal{O}_L \not\models M_n \sqsubseteq C_m$ for any $m > n$. First note that $\mathcal{O}_L \not\models C_n \sqsubseteq \exists s.A$ for any $n \geq 0$ because the same is true for $\mathcal{O}_E$. To obtain $\mathcal{O}_L \models A \sqcap \exists r.C_n \sqsubseteq \exists r.(C_n \sqcap \exists s.A)$, there must exist a subconcept M of $\mathcal{O}_L$ such that $\mathcal{O}_L \models A \sqcap \exists r.C_n \sqsubseteq \exists r.M$ and $\mathcal{O}_L \models M \sqsubseteq C_n \sqcap \exists s.A$. We use M as M_n. Assume to the contrary of what remains to be shown that $\mathcal{O}_L \models M \sqsubseteq C_m$ for some $m > n$. Then $\mathcal{O}_L \models A \sqcap \exists r.C_n \sqsubseteq \exists r.(C_m \sqcap \exists s.A)$, which contradicts the fact that this CI is not entailed by $\mathcal{O}_E$. ❑

Theorem 6. *Let $n \geq 0$ and let $\mathcal{O}_n$ be the union of Γ_n with any of the following sets:*

$$\{\exists r^-.A \sqsubseteq B\}, \quad \{\mathsf{func}(r), A \sqsubseteq A\}, \quad \{r \sqsubseteq s^-, A \sqsubseteq A\}$$

For every $\ell \geq 1$, any ℓ-bounded projective $\mathcal{ELH}$ approximation $\mathcal{O}_L$ of $\mathcal{O}_n$ must be of size at least $\texttt{tower}(\ell, n)$.

Proof. We start with $\mathcal{O}_E = \Gamma_n \cup \{\exists r^-.A \sqsubseteq B\}$. The proof idea is very similar to the proof of Theorem 5. Assume a depth bound $\ell \geq 1$ is given. Take any set Ω of mutually incomparable $\mathcal{EL}(\Sigma_n)$ concepts of depth at most $\ell - 1$ such that

Ω has size $\texttt{tower}(\ell, n)$, where concepts C_1, C_2 are called *incomparable* if neither $\mathcal{O}_E \models C_1 \sqsubseteq C_2$ nor $\mathcal{O}_E \models C_2 \sqsubseteq C_1$. It is straightforward to construct such a set Ω. Then it suffices to show that for every $C \in \Omega$ there exists a subconcept M_C of $\mathcal{O}_L$ such that $\mathcal{O}_L \models M_C \sqsubseteq C$ and $\mathcal{O}_L \not\models M_C \sqsubseteq C'$ for any $C' \in \Omega$ with $C' \neq C$. Assume $C \in \Omega$ is given. Then

$$\mathcal{O}_E \models A \sqcap \exists r.C \sqsubseteq \exists r.(B \sqcap C)$$

Observe that $\mathcal{O}_L \not\models C \sqsubseteq B$. Thus, similarly to the proof above one can show that to obtain $\mathcal{O}_L \models A \sqcap \exists r.C \sqsubseteq \exists r.(B \sqcap C)$ there must exist a subconcept M_C of $\mathcal{O}_L$ such that

- $\mathcal{O}_L \models A \sqcap \exists r.C \sqsubseteq \exists r.M_C$;
- $\mathcal{O}_L \models M_C \sqsubseteq B \sqcap C$.

Observe that $\mathcal{O}_L \not\models M_C \sqsubseteq C'$ for any $C' \in \Omega \setminus \{C\}$ because $\mathcal{O}_E \not\models A \sqcap \exists s.C \sqsubseteq \exists s.(B \sqcap C')$ for any such C'. Thus, M_C is as required.

The proofs for $\mathcal{O}_n = \Gamma_n \cup \{\mathsf{func}(r), A \sqsubseteq A\}$ and $\mathcal{O}'_n = \Gamma_n \cup \{r \sqsubseteq s^-, A \sqsubseteq A\}$ combine the sketch presented for $\Gamma_n \cup \{\exists r^-.A \sqsubseteq B\}$ with the proof idea from Theorem 5. Thus, one considers the same set Ω and then shows that any ℓ-bounded $\mathcal{ELH}$ approximation of $\mathcal{O}_n$ entails all CIs

$$\exists r.A \sqcap \exists r.C \sqsubseteq \exists r.(A \sqcap C)$$

with $C \in \Omega$ and so is of size at least $\texttt{tower}(\ell, n)$, and that any ℓ-bounded $\mathcal{ELH}$ approximation of $\mathcal{O}'_n$ entails all CIs

$$A \sqcap \exists r.C \sqsubseteq \exists r.(C \sqcap \exists s.A)$$

with $C \in \Omega$ and so is of size at least $\texttt{tower}(\ell, n)$. ❑